The Dissertation Committee for Jakob Infuehr

certifies that this is the approved version of the following dissertation:

Relative Performance Evaluation, Analysts, Convertible Debt

Committee:

Volker Laux, Supervisor

Andres Almazan

Michael Clement

Aysa Dordzhieva

Ronghuo Zheng

Relative Performance Evaluation, Analysts, Convertible Debt

by

Jakob Infuehr

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2019

To my parents.

Acknowledgments

I would like to thank my dissertation committee: Andres Almazan, Michael Clement, Aysa Dordzhieva, Volker Laux (chair), and Ronghuo Zheng. Also, I appreciate helpful comments from Prasart Jongjaroenkamol, seminar participants at The University of Texas at Austin, the University of Alberta, the University of Southern Denmark, the Leibniz University of Hanover, the Management Accounting Section Midyear Meeting, and the Tenth Accounting Research Workshop.

Volker Laux, my adviser throughout the Ph.D. program, was of invaluable help. He gave me the freedom and guidance to pursue and achieve my research goals.

I am also indebted to Antonis Kartapanis. His support and friendship the last five years made my life here in Austin a great time. Similarly, Colin Koutney and Xinyu Zhang gave me so much joy. They were always there for me when I needed them most.

Kristen Valentine, Brian Monsen, Zheng Leitter, Ben Van Landuyt, Shannon Garavaglia, Yue Wang, Dan Rimkus, Ryan Hess, Ryan Ballestero, and Skylar DeTure all were very kind to me.

Finally, I thank my parents Maria and Anton Infuehr. I truly appreciate

your continuous support and love. I try to make you proud as best as I can.

JAKOB INFUEHR

The University of Texas at Austin

May 2019

Abstract

Relative Performance Evaluation, Analysts, Convertible Debt

Jakob Infuehr, Ph.D.
The University of Texas at Austin, 2019

Supervisor: Volker Laux

This dissertation addresses several research questions. First, I show that relative performance evaluation incentivizes more earnings management. The optimal contract will depend less on a correlated benchmark, if it is easier for the manager to misreport performance. Thus, the model predicts that firms with strong internal controls and good auditors are more likely to use RPE. Second, I show that a higher analyst following can lead to more or less earnings management, depending on the skewness of the earnings distribution. Analysts try to minimize their forecast error, while managers try to beat the average forecast target. Thus, the actors' actions influence each other. Third, I study the effects of accounting conservatism on the use of convertible debt. In the model, firms use these financial contracts to separate themselves from bad firms, while trying to minimize costs of financial distress. A very aggressive accounting system helps good firms separate and avoid financial distress, because this makes a low signal more informative and less likely, causing investors not to convert to equity after such a signal.

Table of Contents

List of Tables

List of Figures

Chapter 1

Relative Performance Evaluation and Earnings Management

1.1 Introduction

Incentivizing executives to work hard has been studied extensively in the literature. Holmstrom (1982) was one of the first to show that using the performance of the competition in determining compensation can be desirable if there are common shocks that influence output. Filtering this additional noise gives a better understanding of the executive's effort and reduces the risk imposed on the worker. Positive covariance among firms' performances seems to be descriptive of the real world, where industry-wide and economywide events regularly affect multiple firms at once. Hence, one would expect a widespread use of relative performance evaluation (RPE). Yet, empirical research has found only modest use. Bannister and Newman (2003) and Gong, Li, and Shin (2011) find that less than half of firms use RPE.[1] A good example of this is the oil industry. In 2007, at the height of oil prices, executives received (compared to other industries) disproportionately high pay raises, four times as high as the average (Herbst 2008). A similar correlation was observed in 2015. When oil prices plummeted, so did the compensation of many executives. Median compensation fell, while it rose in other industries (Olsen 2016).

Bertrand and Mullainathan (2001) provide one potential explanation for the lack of RPE. They find that "better governed firms pay less for luck," and conclude that this is consistent with the hypothesis that CEOs have power over their boards and essentially set their own pay. In contrast, I show that the fact that weaker corporate governance is associated with less RPE use can also occur in an optimal contracting setting, in which the manager has no power over the compensation committee. An important feature in my model is that the CEO can manipulate the performance measure on which her compensation is based. I find that the optimal contract will make less use of RPE if the CEO can more easily misreport the performance measure.

The intuition for this result is driven by a nonlinearity in the optimal contract. Consistent with empirical evidence (Bannister and Newman 2003,

[1] A current working paper by Bizjak et al. (2018) shows that RPE use has increased to 48% in 2015.

Garvey and Milbourn 2006), in my model the manager receives a large reward for outperforming the benchmark, but her compensation only decreases slightly for underperforming it. This causes two effects, both of which are disadvantages of RPE.

First, the nonlinearity creates a difference in ex-post manipulation incentives. When the manager performs similar to the benchmark, ex-post incentives to manipulate will be higher, while they are lower when she underperforms the benchmark. If the manager's performance were uncorrelated with the benchmark, then these forces would offset, and ex-ante expected manipulation would be unchanged. However, because the manager's performance and the benchmark *are* positively correlated, the chance that the manager performs similar to the benchmark increases. Hence, she is more likely to be in a scenario when she has high incentives to manipulate, and less likely to be in a situation when she has low incentives to manipulate. Without RPE, a contract could still be nonlinear. Yet, the lack of correlation due to the absence of a benchmark means that the manager is *not* more likely to be in a situation when her manipulation incentives are strong. Therefore, RPE in expectation incentivizes a higher level of manipulation compared with a lack of RPE use.

The second effect working against RPE is about the observability of the benchmark. The manager has to commit to her effort decision early in the year, well before the benchmark is realized. Ideally, she would prefer to work hard when the benchmark is easier to beat, and shirk when it is harder.

However, she does not have to commit to an earnings management decision until very late in the fiscal year, possibly even after the fiscal year already ended. Real earnings management decisions such as accelerating next year's sales (by giving discounts) or delaying R&D spending by a month can be made in December; and accrual-based earnings management can be executed even in January, when many major (discretionary) accounting decisions have to be made. At that point, she will have already observed the outcome of shocks, such as changes in economic conditions, industry demand, etc. Thus, the manager can condition the manipulation level on the realization of the benchmark. With nonlinear RPE, her manipulation decision will depend on her performance relative to the benchmark. This makes manipulation more efficient for the agent, hence reducing the incentive effect to actually work hard. Benchmark-independent pay does not have this disadvantage, because the informational advantage is useless to the agent. Her manipulation decision would be independent of economic conditions in that case.

The firm of course anticipates the two above described effects and proactively responds by reducing the weight of relative compensation in the contract structure. It trades off the benefit of using the more informative signal with the benefit of reducing incentives to manipulate. This result can explain why so few firms find it optimal to use RPE, despite the contrarian theoretical prediction. The result also leads to the empirical prediction that RPE will be used more in firms with better corporate governance, tighter accounting standards, etc. that limit the CEO's manipulation potential.[2]

[2]This empirical prediction can be tested differently than the one about powerful CEOs

The remaining paper is organized as follows. Section 1.2 discusses related literature, section 1.3 describes the model, section 1.4 analyzes the scenario without earnings management, section 1.5 discusses the main results and also explores some comparative statics, and section 1.6 concludes. All proofs are in appendix A.

1.2 Related Literature

Relative performance evaluation (RPE) has received some attention in the academic literature. One of the earliest papers is Holmstrom (1982). He showed that the optimal compensation scheme of an agent depends on his performance alone if and only if his performance is independent from anyone else's. In the case of independence, no information about the agent's actions can be learned from comparing his output with another one. It would only add noise and thus be detrimental in the case of risk-aversion because the additional risk that is imposed on the agent has to be compensated. However, with dependence, some information is embedded in the performance of others and this can be used to more precisely incentivize effort. Under some conditions, a weighted average of all performances and the agent's output is a sufficient statistic.

Holmstrom's results did not specify *how* the weighted average should

capturing the pay-setting process. My prediction can use e.g. distance between firm headquarters and the auditor's office (Choi et al. 2012, Kubick et al. 2017), while powerful CEOs can be measured using CEO tenure, CEO-board duality, etc.

be used in the compensation scheme. This is an important question that was addressed in subsequent research. Banker and Datar (1989) find necessary and sufficient conditions when a linear aggregation of signals is optimal. More specifically, using profit is only optimal, when revenue and costs are equally intense (sensitivity times precision) signals. Celentani and Loveira (2006) find that "if the marginal return of effort depends on the aggregate state, optimal contracts are not monotonically decreasing in the performance benchmark" and claim that this may explain the lack of RPE use in the business world. Fleckinger (2012) generalizes these results by not imposing any restrictions on the correlation of outcomes. His results predict that RPE is most effective when covariance is constant and positive.

RPE has, of course, also been studied in the empirical literature. Early papers (e.g. Janakiraman, Lambert, and Larcker 1992, Aggarwal and Samwick 1999) had to resort to indirect tests and did not find positive results. SEC rule changes since then allowed the application of direct methods (e.g. Bannister and Newman 2003, Gong et al. 2011) and results indicated that about a quarter of all firms use RPE. Bannister and Newman (2003) also find, as well as Garvey and Milbourn (2006), that firms use one-sided RPE, executives get rewarded for outperforming the peer-group, but do not get punished for underperformance. This result is consistent with the theoretical results of Celentani and Loveira (2006) and Feriozzi (2011). I also find this in my setting, although the reason is different. In my model, the assumption of limited liability leads to one-sided RPE.

Prior analytical research has proposed solutions for the relative performance evaluation puzzle. Dye (1992) finds that the agent's option to choose among different projects can negatively affect RPE. Gopalan et al. (2010) study a similar idea. The agent will pick a project (or industry) where his skill is relatively high (compared to the peers in that industry), but may be low in absolute terms. This is not in favor of the principal, who cares about absolute profit, not relative one. Another paper is Dikolli et al. (2011), who hypothesize that the empirical literature may have committed Type II errors by summing peer-performance differently compared to boards (selecting more, fewer or different peers). This could lead to a lower estimate on the number of firms that use RPE in compensation contracts.[3]

The only published paper thus far to study the intersection of RPE and earnings management is Bagnoli and Watts (2000). However, their paper does not endogenize the agent's compensation structure and therefore does not examine the effect of earnings management on the use of RPE, key aspects of my paper. There is also no moral hazard problem. A recent working paper by Balakrishnan, Lin, and Sivaramakrishnan (2016) studies ranking systems and its susceptibility to performance manipulation. This paper, though, is more closely related to the tournament theory literature, rather than the RPE literature.

[3]Other papers that discuss the RPE puzzle include Fershtman et al. (2003) and DeMarzo and Kaniel (2017), who analyze agents' utility functions that depend on others' pay.

1.3 Model

I consider a one-shot game with a risk-neutral principal and a risk-neutral agent.

Timing: There are two dates, $t = \{0, 1\}$. At date 0, the principal hires an agent to be in charge of a project and offers her a compensation contract. After signing the contract, the agent can exert costly effort that increases the probability of success. At date 1, the outcome of a correlated benchmark is publicly observed. The agent privately learns whether her project succeeded or failed, $R \in \{S, F\}$.[4] Then, she issues a potentially manipulated report. Based on the report and the benchmark, the agent is paid and the game ends.

Effort: The agent makes a binary effort decision after being hired: $e \in \{e_l, e_h\}$ (shirk or work). The probability of success, $p_e = \Pr(R = S|e)$, is strictly greater when $e = e_h$ is chosen, i.e., $p_h > p_l$. Effort imposes a disutility on the agent, $k(e)$, and without loss of generality, I assume that $k(e_h) = c$ and $k(e_l) = 0$. Furthermore, to avoid trivial solutions, I assume that the principal always wants to incentivize high effort.

Benchmark: There is an observable and contractible benchmark $Z \in \{G, B\}$ (good or bad). The benchmark Z could be the S&P 500, which could be low or high (in this binary model) or other quantitative indices of economic conditions. Commodity prices are another option.[5]

[4] Unobservable outcome might occur, if the project is long-term, but the agent needs to be compensated before the end of the project.

[5] A somewhat more endogenous choice would be the average performance of a peer group. I address this interpretation in section 5.3. However, I deliberately do not use a two agent

$R\backslash Z$	B	G
F	$(1-p_e+\gamma)(1-q)$	$(1-p_e-\gamma)q$
S	$(p_e-\gamma)(1-q)$	$(p_e+\gamma)q$

Table 1.1: Joint Probabilities

The probability that the benchmark is good is q. A pair of results can take four values, $Y \in \{SG, SB, FG, FB\}$. Let the probability of the good benchmark be independent of the agent's effort decision, $\Pr(Z = G|e) = \Pr(Z = G) = q$. However, this does not imply that the probabilities of project success and benchmark are independent. There is a positive covariance γ, which alters the conditional success probabilities such that $\Pr(R = S|Z = G) = p_e + \gamma$ and $\Pr(R = S|Z = B) = p_e - \gamma$. Let γ be small enough such that all probabilities are between zero and one. This leads to the joint probabilities for the four possible outcomes displayed in table 1.1.

These probabilities capture, in the simplest form possible, the character of positive correlation. The probability that the benchmark is good and the agent successful increases. For example, when the economy is doing well, a CEO's projects are more likely to succeed. The assumption of a constant covariance imposes a restriction on the model to keep the focus on the effects of earnings management, and the exposition simple.[6]

Report: At date 1, the agent observes the outcome of her project and the benchmark with certainty. She will then issue a report to the principal,

model. The situation I want to analyze focuses on a CEO, who does not have a comparable coworker within the same company.

[6]The results are robust to alternative specifications.

$r \in \{r_S, r_F\}$. Let the combination of the report and the observable benchmark be denoted by $y \in \{y_{SG}, y_{SB}, y_{FG}, y_{FB}\}$. The agent can take an unobservable, manipulative action $m \in [0, 1]$ to issue a favorable report, even though the project has failed. m is the probability that a failed project will be misreported as a success, i.e. $m = \Pr(r = r_S | R = F)$. The agent would never misreport good news because in the optimal contract compensation is higher when the report is better. Generally, the manipulation level will take two different values, m_G and m_B, depending on the benchmark, good and bad respectively.

The cost of manipulation is $gm^2/2$, where g is an exogenous parameter that captures how easily reports can be manipulated.[7] For example, better auditors, tighter accounting standards, and a vigilant board would cause g to be higher. Furthermore, I assume that $g > \underline{g} \equiv \frac{2c}{p_h - p_l}$ to guarantee an interior solution, i.e. $m < 1$, in equilibrium.[8]

Contracting: The agent is risk-neutral and thus maximizing her expected wage, $E[w_y]$. The only variable that is available for contracting, is the combination of the report and the benchmark, y. Thus, the agent is offered a contract $w = \{w_{SG}, w_{SB}, w_{FG}, w_{FB}\}$ that specifies four state-dependent payments. For example, the agent receives w_{SB} when she reports a success and

[7] One could assume that g depends on the realization of the benchmark. When conditions are bad, auditors may expect more manipulation, thus auditing more. The results of the paper are qualitatively unchanged, as long as the two different g's are close to each other.

[8] These assumptions about the manager's reporting have been used extensively in prior literature with binary models and earnings management (e.g. Dutta and Gigler 2002, Jongjaroenkamol and Laux 2017).

the benchmark is bad, $y = y_{SB}$. The agent is protected by limited liability, that is $w \geq 0$ for each element of the vector. The reservation utility is set to zero. The participation constraint will not bind in equilibrium, and first-best cannot be achieved.

An incentive scheme exhibits RPE, when $\Delta w \equiv (w_{SB} - w_{SG}) > 0$, joint performance evaluation (JPE), when $\Delta w < 0$, and independent performance evaluation (IPE), when $\Delta w = 0$. The greater the difference, the greater the extent of RPE/JPE. w_{FG} and w_{FB} are not part of this definition because in the optimal solution these payments are zero.

1.4 No Manipulation

The no-manipulation setting is a special case of the general model described above, when $g \to \infty$ and therefore $m = 0$. This case demonstrates the main advantage of RPE.

The principal minimizes expected compensation cost

$$\min_{w_{SB},w_{SG},w_{FG},w_{FB}} E[w_y] = (p_h - \gamma)(1-q)w_{SB} + (p_h + \gamma)qw_{SG} + \\ + (1 - p_h - \gamma)qw_{FG} + (1 - p_h + \gamma)(1-q)w_{FB}, \quad (1.1)$$

subject to the agent's incentive constraint

$$
\begin{aligned}
&(p_h-\gamma)(1-q)w_{SB}+(p_h+\gamma)qw_{SG}+(1-p_h-\gamma)qw_{FG}+ \\
&\quad +(1-p_h+\gamma)(1-q)w_{FB}-c \geq (p_l+\gamma)qw_{SG}+(p_l-\gamma)(1-q)w_{SB}+ \\
&\qquad +(1-p_l-\gamma)qw_{FG}+(1-p_l+\gamma)(1-q)w_{FB}, \quad (1.2)
\end{aligned}
$$

and the limited liability constraint

$$w_{SG}, w_{SB}, w_{FG}, w_{FB} \geq 0. \quad (1.3)$$

Solving this problem leads to the following proposition.

Proposition 1.1 When earnings management is not possible ($g \to \infty$), then the optimal contract satisfies

$$w_{SB} > w_{SG} = w_{FG} = w_{FB} = 0. \quad (1.4)$$

It is optimal to set $w_{FG} = w_{FB} = 0$, because any positive payment for failure would only make it more difficult to provide incentives for the agent to work hard. The reason for $w_{SB} > w_{SG}$ is a little more subtle. Holmstrom's work (1979, 1982) shows that it is efficient to incentivize effort by linking pay to the signal that is most informative about effort. If an agent is successful, despite a low benchmark, then success is a very informative signal about the agent's effort. However, high agent performance in the presence of a good benchmark is a less informative signal about effort. Formally, this can be expressed with likelihood ratios:

$$\frac{\Pr(y = y_{SB}|e = e_h)}{\Pr(y = y_{SB}|e = e_l)} > \frac{\Pr(y = y_{SG}|e = e_h)}{\Pr(y = y_{SG}|e = e_l)} \text{ or } \frac{(p_h - \gamma)(1-q)}{(p_l - \gamma)(1-q)} > \frac{(p_h + \gamma)q}{(p_l + \gamma)q}. \tag{1.5}$$

These likelihood ratios measure how strongly $y = y_{SB}$ and $y = y_{ySG}$, respectively, signal that the agent chose high rather than low effort. A high likelihood ratio speaks for high effort; a value of one would indicate that nothing new is learned from the signal (Hart and Holmstrom 1987). The inequality in (1.5) shows that $y = y_{SB}$ (agent is successful and benchmark is bad) is more informative than $y = y_{SG}$.

However, if one were to introduce risk-aversion, then there is a trade-off. A large spread between w_{SB} and w_{SG} imposes a risk on the agent, which has to be compensated in the form of a risk premium. The principal would have to strike a balance between the risk premium and putting more weight on the more informative signal. More risk-averse agents would be offered a contract with less RPE, and agents that have a project with higher covariance would be subject to more RPE (to filter out more correlated noise).

The results in the paper do not rely on risk aversion and the model therefore assumes risk neutrality, to keep the intuition simple. However, all results are robust to the assumption about the curvature of the agent's utility function. Imposing greater risk is a well-known problem associated with RPE. The focus of the paper, though, is on another problem, and risk aversion is thus ignored.

1.5 Manipulation

Consider now the setting of main interest, in which earnings management is possible, that is $g < \infty$.

1.5.1 Exogenous Contract

I first analyze the manager's optimal level of manipulation for any exogenous contract which satisfies the incentive constraint with equality. I set $w_{FB} = w_{FG} = 0$, as these payments do not incentivize high effort. Taking the derivative of the agent's utility function (A.13) with respect to m_B and m_G, respectively, yields

$$m_B = \frac{w_{SB}}{g} \text{ and } m_G = \frac{w_{SG}}{g}. \tag{1.6}$$

Manipulation depends on the outcome of the benchmark, Z, because the agent can observe Z before the manipulation decision. In contrast, effort cannot be conditioned on Z because the manager only observes Z after the effort decision.

The expected manipulation level $E[m]$, in general terms, is

$$E[m] = (1 - p_e + \gamma)(1 - q)m_B + (1 - p_e - \gamma)qm_G. \tag{1.7}$$

For the expected manipulation level, the following proposition can be obtained.

Proposition 1.2 Expected manipulation level $E[m]$ is always higher for RPE ($w_{SB} > w_{SG}$) than for benchmark-independent compensation ($w_{SB} = w_{SG}$). $E[m]$ is minimized, when JPE is used. The minimal manipulation inducing contract is characterized by an interior solution $w_{SG} > w_{SB} > 0$, when earnings management is easy ($g < g_T$), and by the corner solution $w_{SG} > w_{SB} = 0$, when $g > g_T$.[9]

Solving for the contract that satisfies the incentive constraint (A.16) with equality, plugging all results into (1.7), taking the derivative of $E[m]$ with respect to w_{SB}, and substituting w_{SG} back in for simplicity, gives the following result:[10]

$$\frac{dE[m]}{dw_{SB}} = \frac{1}{g}(1 - p_h + \gamma)(1 - q) - \frac{1}{g}(1 - p_h - \gamma)\frac{g - w_{SB}}{g - w_{SG}}(1 - q). \qquad (1.8)$$

The first summand in (1.8) is the direct effect of an increase in w_{SB} on m_B. The manager will manipulate more when the benchmark is low (m_B), if the bonus payment he receives when he is successful and the benchmark is low, w_{SB}, is higher. The second summand in (1.8) is the indirect effect of an accompanied decrease in w_{SG} and hence m_G. Several insights can be gained from (1.8). First, when $w_{SB} > w_{SG}$ (RPE), then the derivative in (1.8) is positive, i.e. an increase in RPE always increases expected manipulation.

[9] g_T is defined in the appendix, equation (A.20).

[10] Note that in a contract that satisfies the incentive constraint with equality, w_{SB} and w_{SG} are negatively related. When one payment increases, the other can be lowered. Thus, a change in w_{SB} is always in the same direction as a change in $\Delta w \equiv (w_{SB} - w_{SG})$.

This result can be traced back to the positive correlation between the manager's performance and the benchmark, as well as the nonlinearity of the contract. Nonlinear RPE creates incentives to manipulate more when the benchmark is bad, because $w_{SB} - w_{FB} > w_{SG} - w_{FG}$. Manipulation allows the manager to report that she outperformed the benchmark and collect a big bonus. This scenario is more likely to occur when positive correlation is at play because the agent is likely to perform similar to the benchmark. Additionally, the optimal contract creates lower incentives to manipulate when the benchmark is good. Manipulation only allows the CEO to hide her underperformance relative to the benchmark, and avoid a small penalty (equivalently: receive a small bonus). This case is less likely to occur due to correlation because the agent's performance is less likely to deviate from the benchmark. Since these forces flip when joint performance evaluation (JPE) is used, $w_{SB} < w_{SG}$, expected manipulation is lower with JPE.

The fact that expected manipulation can have an interior minimum, when manipulation is cheap, is driven by another key force in the model. That force concerns the observability of the benchmark. If it is observable before the manipulation decision, then the manipulation level, when the benchmark is good, will differ from the level, when it is bad (as long as $w_{SB} \neq w_{SG}$). If it were unobservable, then by construction the agent would have to choose $m_B = m_G$. The loss of a degree of freedom for the agent means that she will receive a lower utility to the benefit of the principal. He can pay out the larger bonus less often because the agent is less efficient in her manipulation decision.

This leads to the following proposition.

Proposition 1.3 If compensation depends on the benchmark ($w_{SB} \neq w_{SG}$), then the agent's utility is higher when she can observe the benchmark.

This result is a drawback for both RPE and JPE, and explains why expected manipulation need not be monotonically decreasing as JPE usage increases.

This second effect also is true due to the asymmetric contract induced by limited liability. When the agent submits a low report, then compensation is zero, no matter what. If compensation were symmetric, such that $w_{SB} - w_{SG} = w_{FB} - w_{FG}$, then manipulation incentives would be independent of the benchmark. This is an important point to emphasize. In a LEN model, the linear restriction imposed on the contract makes it symmetric and thus this effect would not be present. The asymmetric contract can arise in many situations, depending for example on how effort affects covariance or limited liability (as is the case in this paper), and is consistent with evidence from archival studies (e.g. Bannister and Newman 2003).

1.5.2 Endogenous Contract

I now turn to the endogenous contract setting when manipulation is possible. It is essentially the patterns described in Proposition 1.2 and Proposition 1.3 that lead to the following proposition describing the optimal contract.

Proposition 1.4 The optimal contract has the following characteristics:

(i) $w_{SB} > w_{SG} = 0$ (max RPE) when $g \in [g^*, \infty)$,[11]

(ii) $w_{SB} > w_{SG} > 0$ (RPE) when $g \in (\widehat{g}, g^*)$,

(iii) $w_{SG} = w_{SB}$ (IPE) when $g = \widehat{g} \equiv \frac{8c}{3(p_h - p_l)}$,

(iv) $w_{SG} > w_{SB} > 0$ (JPE) when $g \in (\underline{g}, \widehat{g})$,

(v) $w_{SG} - w_{SB} \to 0$ (*asymptotically IPE*) when $g \to \underline{g} \equiv \frac{2c}{p_h - p_l}$.

The possibility of report manipulation creates a trade-off in the RPE usage. On the one hand, RPE is beneficial due to the positive correlation (as seen in the benchmark scenario). On the other hand, RPE incentivizes higher manipulation. As a result, when the use of RPE indeed causes a significantly higher manipulation level, then the principal will be more inclined to abstain from benchmarking compensation. In fact, this effect can be high enough, such that the principal actually prefers joint performance evaluation (JPE), i.e. $w_{SB} < w_{SG}$, for some levels of g that are close enough to $\underline{g}$. This means that the agent actually earns a higher wage, when the competition was also successful. The fact that the agent can observe the benchmark before her manipulation decision is also a disadvantage for JPE and thus favors independent compensation. However, the other key intuition that worked against RPE, works in favor of JPE. The agent wants to report a success while the benchmark is high. This outcome is more likely due to the positive correlation, thus effort is more likely to achieve the desired result. On the other hand, the lack of correlation for manipulation does not help it, and thus, it is not as efficient in reaching the agent's preferred outcome. Of course, JPE runs

[11] g^* is defined in the appendix, equations (A.25) and (A.26).

counter to using the most informative signal, but when the manager can easily manipulate (g close to $\underline{g}$), then its advantages can outweigh the drawbacks. At $g = \widehat{g}$, the forces balance exactly, and independent performance evaluation is optimal.

Part (v) of Proposition 1.4 follows because the principal has to take manipulation into account, and a very uneven compensation structure incentivizes heavy manipulation because the agent can wait and observe the outcome of the benchmark. For example, for $w_{SB} > 0$ and $w_{SG} = 0$, she would manipulate heavily when the benchmark is low and not at all when the benchmark is high. Given the convex manipulation cost function, the agent has to be compensated for this manipulation (or else she shirks) with a very high payoff that can be avoided by smoothing compensation across states. As a result, in addition to the forces described in the previous section, RPE now has an additional disadvantage when the agent can wait and observe the outcome of the benchmark. The agent is now in a better position to efficiently exploit RPE because she will not waste her manipulation effort when it is not as beneficial to her.

Figure 1.1 illustrates an example of the above described observations.[12]

In the figure, $g^* \approx 173$. It is the point where w_{SG} hits the x-axis. For $g \geq g^*$ (not pictured), when manipulation is very costly, the agent only receives money if she reports success and the benchmark is low. The graph shows the main relationship. As earnings management becomes less costly

[12]The example uses the parameters $p_h = 0.55$, $p_l = 0.45$, $\gamma = 0.4$, $c = 2$, and $q = 0.55$.

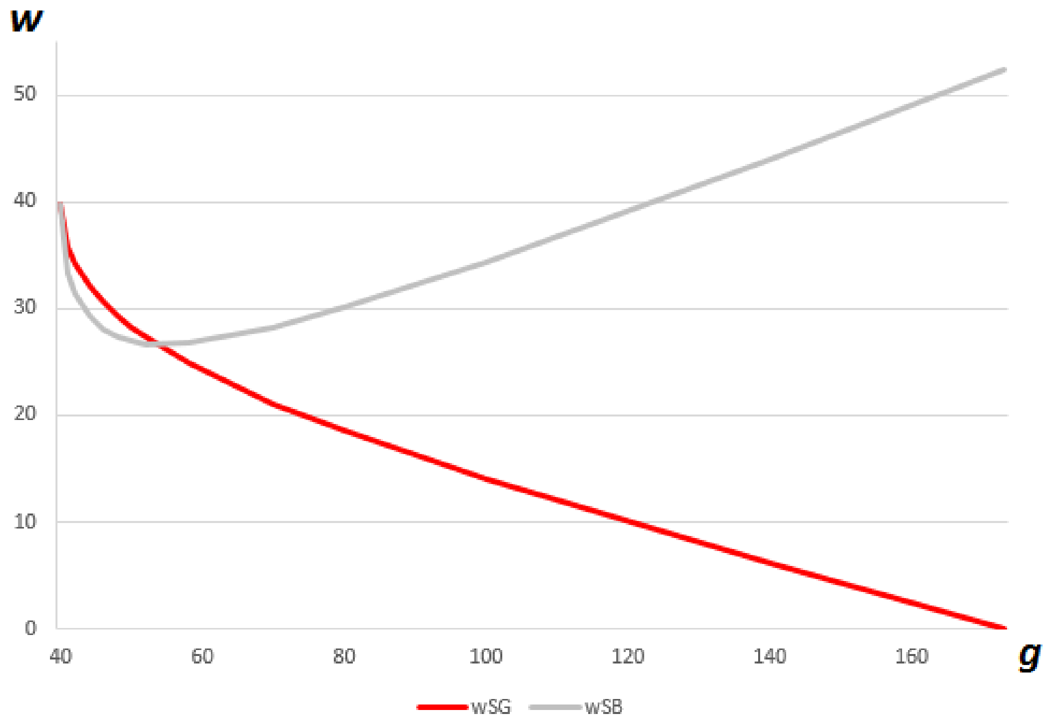

Figure 1.1: Optimal Compensation

(moving from right to left on the graph), the firm has to reduce the importance of the benchmark by decreasing the difference between the two wage payments w_{SB} and w_{SG}. At the very left end of the chart (when g approaches $\underline{g} = 40$), incentivizing the agent to work hard becomes tough enough that both wage payments have to increase to accomplish this feat. For $g < \underline{g}$, no solution exists. Threre is also a small range, $g \in (40, \frac{160}{3})$, where joint performance evaluation is optimal, i.e. $w_{SG} > w_{SB}$. At $g = \frac{160}{3}$, the two wage payments are identical.

1.5.3 Comparative Statics

I now turn the analysis to some comparative statics. Less RPE is used when either the range "max RPE" in part (i) of Proposition 4 shrinks (i.e. g^* increases) or the range "RPE" in part (ii) of Proposition 4 gets smaller (i.e. $\widehat{g}$ increases).

Corrolary 1.1 RPE is used less as

(i) covariance (γ) decreases,

(ii) probability of success (p_h) decreases, and

(iii) cost of effort (c) increases.

A change in correlation has two effects. First, when correlation increases, then this increases the advantage of RPE as seen in the proof of the no manipulation scenario (see appendix). The likelihood ratios diverge and it is increasingly easier to incentivize effort using RPE. Second, an increase in correlation increases the difference in manipulation levels that RPE and independent compensation cause. While earnings management does not change if benchmark-independent pay is used, it does with RPE. This is again owed to the fact that an increase in correlation makes it harder to outperform the competition by regular means, which is why the promised compensation has to increase. This increased compensation in turn causes the increase in manipulation. However, the second effect is lower than the main effect on the informativeness of the signal.

An increase in p_h is beneficial for RPE because it alleviates the incentive problem. The agent is more likely to succeed if he puts in the effort, and effort matters more because p_l is kept constant, increasing $(p_h - p_l)$. Thus, the principal does not need to offer as much compensation which of course also decreases the manipulation incentives. And as already established, a principal that is less concerned about manipulation, will use RPE more heavily. The effect of c can best be understood, when the ratio $\frac{c}{g}$ is viewed as the relative cost of effort in relation to the cost of manipulation. Characteristics like optimal use of RPE and equilibrium manipulation level are unaffected, as long as $\frac{c}{g}$ stays constant. It is thus not surprising that an increase in c causes a proportional increase in g^*.

Corollary 1.2 As manipulation becomes more difficult (g increases),

(i) expected manipulation goes down,

(ii) expected wage decreases,

(iii) the agent's utility decreases.

Part (i) of the corollary can be explained as follows. When manipulation is more costly to the agent, holding everything else equal, he will manipulate less. A counteracting force here is the fact that with less manipulation, the principal will use more RPE. However, this force is weaker than the first one. If the principal would increase RPE too much and cause more, not less, manipulation, then he would again suffer too much from the negative consequences of RPE, and compensation would not be optimal. Hence, RPE will only increase slowly and expected manipulation will decrease.

Part (ii) of the corollary follows because less manipulation makes it easier to incentivize effort. As noted above, the ratio $\frac{c}{g}$ is important as the agent always weighs her options to achieve the desired result. So when manipulation is harder, the bonus necessary to induce effort declines.

Part (iii) follows because even though expected wage decreases, it is not immediately obvious that this implies that the agent's utility also decreases. After all, manipulation decreases and consequently also the cost of manipulation. Why is it not possible that the second effect outweighs the first one? It is instructive to note that the agent can freely choose the manipulation level. When g decreases, she has the option to maintain the original manipulation levels, m_B and m_G. This implies that she could keep the manipulation cost down while simply enjoying the increased wage. The only reason why the agent increases manipulation, is because it increases her overall utility. There is no reason why the agent would intentionally hurt herself and cause her utility to decline.

1.5.4 A manipulated benchmark

Thus far I have assumed that the benchmark is completely exogenous. However, a reasonable argument can be made that executives at the other companies can manage earnings as well. In this subsection I analyze how the optimal contract and the agent's behavior changes in response to a change in q, the probability that the benchmark is good. While this does not endogenize the benchmark completely, it is a reasonable approximation in situations when each individual firm is only a small part of the peer-group. When that is the case, each individual manager's actions will have negligible effects on other firms' decisions. A higher q can be a stand-in for other firms also manipulating their earnings, thus causing the benchmark to be higher more frequently.

The most direct and obvious effect of an increase in q (when other firms manipulate more) is that it becomes harder to outperform the benchmark. The probability that the agent is successful while the benchmark is low decreases. This in turn requires an increased bonus, if one wants to use RPE, to compensate for the decreased likelihood of achieving the bonus. That bigger bonus, however, incentivizes more manipulation, and the board then uses less RPE.

Corollary 1.3 RPE is used less (g^* increases) as the probability of a high benchmark, q, increases.

1.5.5 Renegotiation

One aspect that often receives some attention in the literature is renegotiation. On the one hand, allowing for renegotiation undermines commitment. On the other hand, it can also be used to incorporate new information into the contract that was not available at the beginning.

To illustrate the former point, imagine that in the present model renegotiation is possible after the agent has made her effort decision. If the agent is risk-averse, then renegotiation can be mutually beneficial, because the agent bears some risk that can be completely transferred to the principal. The agent receives a flat wage, regardless of outcome. Unfortunately, this can and will be anticipated by the agent. Realizing that, ultimately, compensation will be independent of outcome, and the agent has no incentive to work.

Illustrating the latter point is possible if renegotiation is possible after the agent privately observes the outcome of her project. At first, it might seem that the principal is in a bad spot to renegotiate because the agent has an informational advantage about the project outcome. However, this does not really matter because if the agent was successful, then she will not accept any new contract that would reduce the pay that she is guaranteed to receive. After failure, renegotiation can prevent earnings management because the agent will receive a flat wage. This time, if the principal has all the bargaining power, it does not affect ex-ante effort incentives. This is because expected ex-ante utilities, depending on outcome, do not change. The agent is still interested in achieving success because his pay will be higher in that case. However, because earnings management will not occur, this can be anticipated by the principal and will therefore affect the initial contract that is offered to the agent. The contract that would be offered without renegotiation would still satisfy the agent's incentive constraint. The key difference is that the principal can save the agent's manipulation cost via renegotation. Hence, a contract that would lead to more manipulation, if renegotation were not available, can be offered. Thus, the optimal contract will feature a heavier reliance on RPE, since RPE again would incentivize more earnings management.

Proposition 1.5 If renegotiation is possible after the agent observed the outcome and the principal has all the bargaining power, then there exists a new threshold g_n^* such that the optimal contract sets $w_{SB} > 0$ and $w_{SG} = 0$

if and only if $g \geq g_n^*$. The new threshold is lower than the old one, $g_n^* < g^*$, i.e. more RPE is used. Additionally, JPE is never optimal.

1.6 Conclusion

This paper analyzes the effect of earnings management on the profitability of relative performance evaluation (RPE). When the manager cannot misreport, using RRP is optimal because it filters industry-wide shocks and allows for a more accurate measure of the CEO's performance. High firm performance when the competition failed is a more informative signal than good performance when others succeeded, too. This insight would predict that RPE should be widespread in the business world, given that positive correlations are common. However, empirical research shows that this is not the case. I show that one potential explanation for this can be earnings management. Manipulation makes it easier for a manager to report success, while the competitors fail. This is the case because manipulation success is less correlated than real outcomes. Whether an auditor for company A demands an adjustment will depend less on whether another auditor demanded an adjustment for company B. Additionally, when the manager can observe the benchmark before her earnings management decision, then RPE increases manipulation incentives even further because the agent will manipulate more when it is more advantageous to her. This gives the CEO an efficient way to boost her expected compensation and thus makes it harder to incentivize hard work.

Ultimately, the firm will anticipate the manager's incentives to manipulate and will switch to a compensation scheme that depends less on the competition in an effort to curb manipulation. Hence, the firm trades off the benefit of using the most informative signal with the drawback of incentivizing manipulation. This result can explain why RPE is not as heavily used as predicted. For some companies, the benefit of RPE will not be big enough compared to its disadvantage.

Chapter 2

Meet or beat analysts' forecasts, skewness, earnings management, and investment choice

2.1 Introduction

Analysts, in their role as information intermediaries, affect firms directly and indirectly in many ways. One of these is corporate governance. Yu (2008) and Chen et al. (2015) have found that a higher analyst coverage improves corporate governance, first and foremost by reducing earnings management. This finding lends support to the monitoring hypothesis: analysts study financial statements thoroughly, take part in conference calls with top level executives, and increase attention due to the dissemination of information.

Contrary to these results, there is recent research that finds negative consequences of increased analyst coverage. He and Tian (2013) find that firms with more analysts are less innovative, and Huang et al. (2017) show that more analyst coverage can cause firms to meet the forecast target more often (which implies more, not less, manipulation). These results are taken as support for the pressure theory, which claims that analysts create pressure on managers to beat the forecast target. This theory is somewhat in conflict with the monitoring theory. However, there is another channel that indirectly affects manipulation decisions. Managers benefit from meeting or beating the analyst forecast target (Bartov et al. 2002) and hence make discretionary accounting choices to accomplish just that (e.g. Matsumoto 2002, Brown 2001). In this paper, I show that the number of analysts following a firm affects this channel. Hence, the observed empirical patterns can be explained, even when there is no analyst monitoring or pressure at all.

Specifically, I consider a model in which managers will manipulate earnings to just meet the average analyst forecast, if actual earnings are close enough to the target. I call this the "manipulation range" and it shows in a histogram as a discontinuity gap. Analysts, on the other hand, try to minimize the forecast error between their (individual) forecasts and reported earnings (actual earnings are unobservable). Hence, managers are influenced by analysts' forecasts, and analysts are influenced by managers' choice of manipulation. I analyze two scenarios: when there are many analysts, and when there is only one covering the firm.[1] In the former case, each analyst takes the

[1]The intuition gained from these two extreme cases applies to intermediate situations as

anticipated reported earnings distribution as given, and tries to minimize his or her forecast error accordingly. An equilibrium will form, when the analysts' forecasts induce a reporting strategy by the manager, such that no analyst can change his forecast without increasing the forecast error.

This equilibrium will change, if there is only one analyst issuing an earnings forecast. Now, the lone analyst anticipates that a change in his forecast also changes the reported earnings distribution. That creates an incentive to decrease the forecast (compared to the equilibrium desribed above). If the reported earnings distribution wouldn't change, then a small decrease in the forecast would not increase the forecast error by much (the first derivative of the forecast error at the optimum is zero). Also, if actual earnings fall in a range, where the manager would manipulate to the target either way, then this also does not affect the error. The change in error will occur on the edges of the manipulation range, especially the low end. When the manager would not manipulate in the original equilibrium because earnings were a little too far away from the target, he would manipulate now with the lowered target. This decreases the forecast error for this specific earnings realization to zero. On the high end, when actual earnings are just below or above the target, the change is negligible, because either way the forecast error will be small. In less technical words, the manipulation range is advantageous to the analyst because it reduces his forecast error. By reducing the forecast, the manager will manipulate slightly lower earnings that used to be farther away from the

well. The number of analysts ranges widely in the real world.

forecast target.

The effect of a lower forecast target on the average accounting manipulation depends on the shape of the earnings distribution. When there are many analysts, then the forecast that minimizes the squared error is the mean of reported earnings (or median for absolute error, Gu and Wu 2003). When there is only one analyst, a lower forecast causes more manipulation if the earnings distribution (probability density function) at the mean is decreasing. The manipulation range shifts left, and thus actual earnings are more likely to fall into that range. This is the case when the earnings distribution is positively skewed as then the mode of the distribution is lower than the mean. The reverse is true for a left-skewed distribution. Actual earnings are less likely to fall into the manipulation range when the number of analysts covering the firm decreases.

This result shows that analyst following can impact earnings management, even when analysts do not monitor the firm, when earnings are skewed. Earnings skewness is an overlooked attribute in prior studies (Yu 2008, Chen et al. 2015) on this topic. Gu and Wu (2003) find that the median skewness is about zero, but there is "large cross-sectional variation in earnings skewness." Hence, it seems plausible that average skewness in other studies could be positive or negative, depending on the specific sample that is used. Therefore, I propose that skewness should be a control variable in future research to clearly identify the forces driving the relationship.

Managers can affect the reported earnings distribution not only by changing the reporting strategy, but also by changing the underlying actual earnings distribution via investments. A key difference between the two is that investment decisions are usually long-term, and thus have to be made well in advance of analysts issuing their forecasts. Therefore, analysts will take the investment decision as given, regardless of the number of analysts, and no differing equilibria emerge. The manager still cares about beating the average forecast as often as possible. If the earnings distribution is symmetric, then there is no incentive to deviate from first-best, because first-best gives the manager a 50% chance to beat the forecast, which he cannot improve upon. However, when the distribution is skewed, and if analysts forecast the median, then a deviation from first-best is optimal for the executive. First-best investment will maximize the expected value, while the manager cares about the median. If median and mean are different (due to skewness), then he has an incentive to improve the median, even when this comes at the expense of a decrease in the mean. Depending on his opportunity set, this could lead to a preference for right-skewed earnings distributions.

The above result shows that it can be perfectly rational for a manager to prefer a more right-skewed earnings distribution, as observed by Schneider and Spalt (2016). They find that "capital expenditure is increasing in the expected skewness of segment returns." However, they attribute this result to the fact that managers use gut feel, which causes them to subjectively assess probabilities, overweighting low probability events, as predicted by prospect theory (Kahnemann and Tversky 1979). Instead, I show that managers' desire to beat the analyst forecast can be the driving force behind this observed pattern.

Finally, I show that the interaction of investment and reporting decisions incentivizes the manager to underinvest (i.e. reduce variance) in order to improve the probability that earnings management allows him to beat the target. Thus, for left-skewed distributions, the possibility of earnings management can improve investment efficiency. The incentive to increase variance in order to improve the median, and the incentive to decrease variance to improve the efficiency of accounting manipulation can cancel each other out.

The remaining paper is organized as follows. Section 2.2 discusses related literature, section 2.3 describes the model, section 2.4 analyzes the scenario when earnings management is possible, section 2.5 discusses the case when the manager's investment choice is endogenous, section 2.6 examines the interaction between manipulation and investment, and section 2.7 concludes. All proofs are in appendix B.

2.2 Related Literature

There are several streams of literature relevant for this paper. The first one is the literatue dealing with managers' incentives to meet or beat the analyst forecast target. Brown (2001) finds that during the 1990s, median forecast errors have shifted from small negative to small positive, indicating an increased incentive to meet or beat the average forecast. Additionally, the share of small positive surprises increased over time, and is more pronounced in growth firms. Brown and Caylor (2005) confirm this trend, and further show that since the mid-1990s the managerial tendency to avoid negative earnings surprises is stronger than the one to avoid negative earnings or earnings decreases. They attribute this pattern to "increased media coverage given to analyst forecasts, more analyst following, more firms covered by analysts, and temporal increases in both the accuracy and precision of analyst forecasts." Furthermore, Armstrong et al. (2017) find that the analysts' external EPS goal is more important to managers than internal incentive plan EPS goals.

Additional research tried to pinpoint the exact source of managerial incentives. Mutsunaga and Park (2001) examine the CEO's annual discretionary bonus (allocated from a bonus pool) and find that the bonus is negatively affected if the consensus analyst forecast is not met for at least two quarters of the year, while there is no significant effect for loss quarters. This reduction in bonus pay is in addition to the traditional linear pay-for-performance sensitivity. Bartov et al. (2002) study the stock market's reaction to barely beating the analysts' earnings expectations. They find that there exists a market premium for beating expectations, even "after controlling for the earnings forecast error for that period." This premium exists, even when beating expectations was probably achieved via earnings management. Furthermore, the premium does not fade over a longer time window, indicating that "investors rationally react to the earnings surprises." Finally, Matsumoto (2002) looks at firm characteristics that are associated with an increased likelihood to beat expectations. These include "firms with higher transient institutional ownership, greater reliance on implicit claims with their stakeholders, and higher value-relevance of earning." These characteristics may by related to increased incentives for managers to beat expectations. Similar to other studies, Matsumoto finds that managers, in addition to accounting manipulation, guide analysts' expectations downward, to increase chances of beating these.

The second stream of literature is the analyst following literature. There are of course many firm characteristics that determine analyst following (e.g. Bhushan 1989: size, institutional ownership, return volatility, market return, business complexity), but the one particularly related to this paper is about earnings management. Previts et al. (1994) in their qualitative study find that analysts are aware of "adjustments of conservative, discretionary reserves, allowances, and off-balance-sheet assets" and noted that analysts seemed to like these adjustments. This observation is consistent with my theoretical model, which finds that analysts benefit from earnings management due to the reduction in forecast error.

More relevant than the determinants of analyst coverage, which are exogenous in my model, are studies about the effect of analyst following on firms. Quite a few papers find positive effects of analyst coverage on firms, such as market value (Chung and Jo 1996), less earnings management (Yu 2008), less excess CEO compensation, and pay-for-performance sensitivity (Chen et al. 2015). All of these studies attribute these positive effects to the analyst monitoring hypothesis, which states that analysts apply pressure on management by increasing investor attention and disseminating information. On the other hand, there is recent research that finds negative consequences of increased analyst coverage. He and Tian (2013) find that firms with more analysts are less innovative, and Huang et al. (2017) show that more analyst coverage can cause firms to meet the forecast target more often. These results are taken as support for the pressure theory, which claims that analysts create pressure on managers to beat the forecast target. This theory is somewhat in conflict with the monitoring theory. Contrary to these theories, I posit and show analytically that the effects can also possibly be explained without the need for monitoring or pressure.

The final stream of literature is related to earnings distribution skewness. Skewness causes mean and median to diverge, which is relevant when making an assumption about analysts' utility function. Gu and Wu (2003) and Basu and Markov (2004) make the case that analysts minimize their absolute forecast error, and hence forecast the median. Lastly, Schneider and Spalt (2016) study the consequences of earnings skewness on investment decisions. They find that CEOs overinvest, when earnings distributions are right-skewed, consistent with the long-shot bias proposed by the prospect theory (Kahneman and Tversky 1979). In this paper, I show that this observed pattern can be rationally explained by management's desire to outperform earnings expectations.

2.3 Model

I consider a one-period game with three dates, a firm manager and one or more analysts.

Timing: There are three dates, $t = \{-1, 0, 1\}$. At $t = -1$, the manager invests k into a project that yields a return at $t = 1$. At $t = 0$, each analyst simultaneously issues a forecast, predicting firm earnings. At $t = 1$, earnings realize, but are not observable to anyone but the firm manager. Then, the manager issues a potentially manipulated report, each player receives their payoff, and the game ends.

Earnings and investment: At $t = 1$, unobservable earnings x are realized. The probability density function is $f(x)$, the cumulative distribution function is $F(x)$, the mean and variance depend on the investment, $\frac{d^2E[x]}{dk^2} < 0$, there exists $k = k_{FB}$, such that $\frac{dE[x]}{dk} = 0$, $Var[x] > 0$ for all k, and $\frac{dVar[x]}{dk} > 0$ for all k. While the location and scale parameters of the distribution depend on investment, the shape parameters (skewness, kurtosis) do not, they are constant. For ease of exposition, I assume that the distribution is single-peaked. The assumptions about the earnings distribution capture the notion that any deviation from the optimal amount of risk, will decrease expected earnings. Also, I will examine both observable and unobservable investment.

Manager and report: I assume that the manager's only incentive is to meet or beat the mean analyst earnings forecast z:

$$\max \int_z^\infty g(y)dy, \tag{2.1}$$

where $g(y)$ is the probability density function of reported earnings y as defined below.

The manipulation results change slightly, if the median forecast matters, and this will be addressed. The investment results are robust. If earnings x are sufficiently close to the forecast z $(0 < z - x < m)$, then he will issue a manipulated report y, and otherwise issue an unmanipulated report:

$$y = \left\{ \begin{array}{c} z \text{ if } z - m \leq x \leq z \\ x \text{ otherwise} \end{array} \right\}. \tag{2.2}$$

Also, I will refer to $[z - m, z]$ as the "manipulation range" because the manager will only manipulate if earnings fall within this range. The assumptions above capture, in the simplest form possible, the empirically observed fact that managers have incentives to meet or beat the analysts' forecast target. Figure 2.1 illustrates a possible example.

Analysts: The analysts know the earnings distribution, $f(x)$. Each analyst's goal is to minimize the squared difference between his or her forecast and reported earnings, i.e. $\min_z E[(z - y)^2]$. They are all equally well-informed and thus, in equilibrium, will all issue the same forecast z. These assumptions warrant some discussion. The key characteristic of the utility function is that the analyst suffers a loss that is increasing in the forecast error. Alternative specifications are possible, such as minimizing the absolute error or maximizing expected reputation (if there were multiple analyst types). The results for investment depend on this assumption, and will receive further discussion

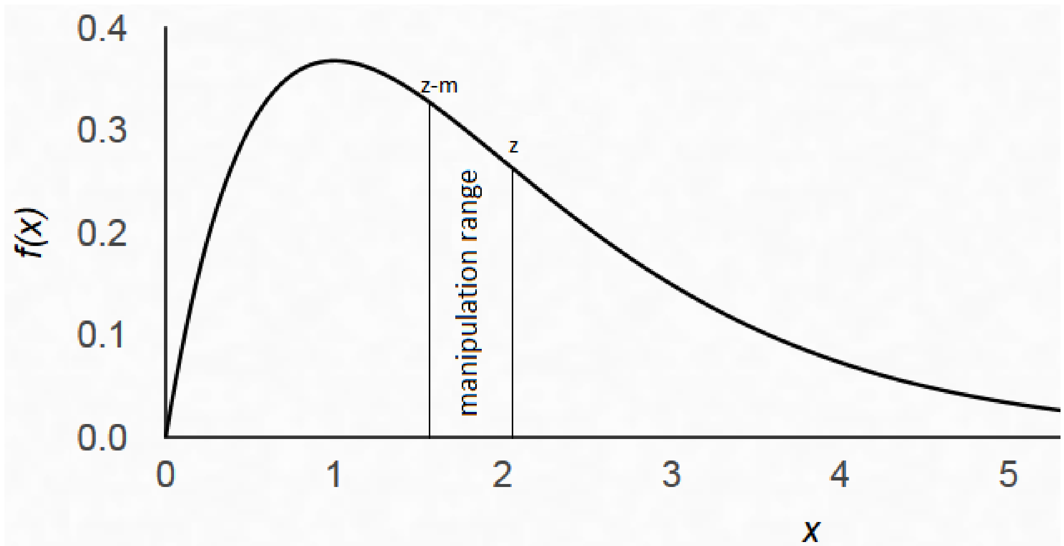

Figure 2.1: In this example, unmanipulated earnings x are distributed in the form of a Gamma distribution: $f(x) \sim \Gamma(2,1)$. The forecast is $z = 2.05$. The manipulation range width is $m = 0.5$. Thus, the manipulation range is $[1.55, 2.05]$. For these values, the mean of reported earnings will be equal to the forecast, an equilibrium, when there are many analysts.

below. Additionally, there is no forecast dispersion in this model, because all analysts are equally well informed. This assumption simplifies the exposition, but does not affect any inferences.

2.4 Earnings management

In this section, I assume that investment is exogenously given, and normalize $E[x] = 0$. Thus, the complete focus of this section is on the effects of earnings management. Furthermore, for mathematical convenience, I assume that m is small compared to $Var[x]$, i.e. $\frac{m}{Var[x]} \to 0$.

2.4.1 Many Analysts

In this subsection, I assume that there are infinitely many analysts. Each analyst's effect on the average forecast target is negligible. As a result, every analyst takes the manager's reporting strategy as given. The analysts minimize their squared error, which means that their forecast has to be the expected value of the reported earnings distribution:

$$z = E[y] = \int_{-\infty}^{z-m} xf(x)dx + z(F(z) - F(z-m)) + \int_{z}^{\infty} xf(x)dx. \tag{2.3}$$

The first and last summands are the unmanipulated reports, when earnings are either too far away from the target or higher than the target and no manipulation is necessary. The middle term represents reported earnings when actual earnings are just a hair below the average forecast. In that instance, the manager would manipulate the report to meet the forecast. Since reported earnings are higher than actual unobservable earnings by the amount of manipulation, the above equation can also be written as:

$$z = E[y] = \underbrace{E[x]}_{=0} + \int_{z-m}^{z} (z-x)f(x)dx. \tag{2.4}$$

Thus, equation (2.4) shows expected manipulation, when there are many analysts. Note that they are evaluated based on the difference between their forecast and *reported* earnings. Analysts have to anticipate the manager's expected amount of manipulation and adjust their forecasts upward accordingly, otherwise their forecasts would be too low. In equilibrium, the mean of the reported earnings distribution influenced by the analysts' forecasts has to match the analysts' mean forecast. Only then will analysts not have any incentive to deviate from the equilibrium strategy. I do not need to solve for z explicitly, since I am not interested in the location of z, but how z changes as the number of analysts changes.

2.4.2 One Analyst

I now analyze the situation, when there is only one analyst. The analyst will take the effect of his report on the reporting strategy of the manager into account, and due to this indirect effect, the forecast will deviate from the expected value of reported earnings. The analyst's objective is as follows:

$$\min_z E[(z-y)^2]. \tag{2.5}$$

Since I only need to find out whether the forecast will be lower or higher than the one in the last section, it is sufficient to analyze the change in squared error at the expected value. At that value, the direct effect will be zero, because the derivative at an optimum is zero. Small deviations from an optimum do not matter much. Therefore, it will be the indirect effect via the change in the manager's reporting strategy that determines the new equilibrium. Let the "many analysts equilibrium forecast" be z_m and the "one analyst equilibrium forecast" be z_o. Assume that $z_o < z_m$ (to be verified later). Then there are two ranges, where the reporting strategy changes: at the lower end of the manipulation range $[z_o - m, z_m - m]$ and at the higher end of the manipulation range $[z_o, z_m]$. Figure 2.2 illustrates this.

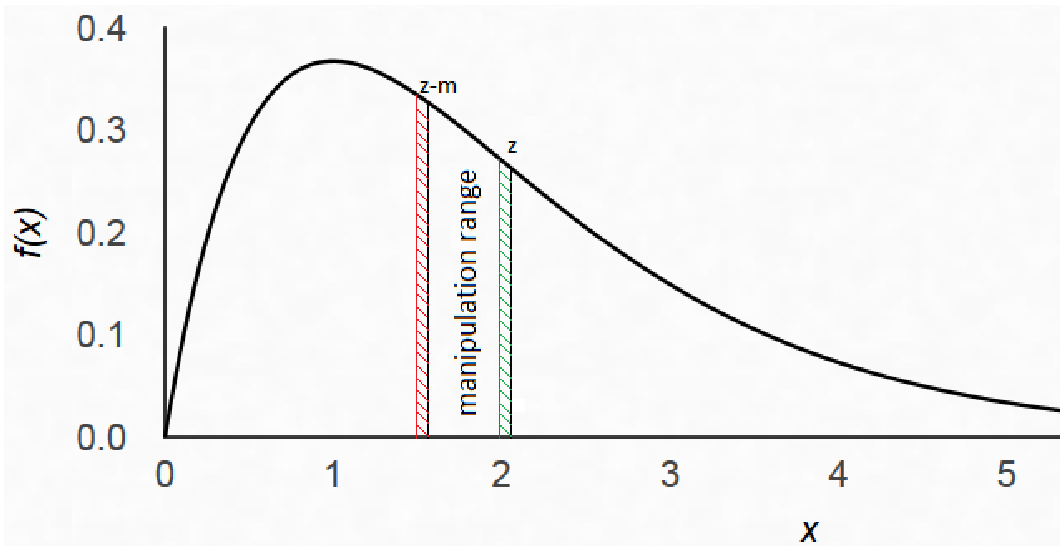

Figure 2.2: This is the same distribution as in Figure 2.1: $f(x) \sim \Gamma(2,1)$. When there are few analysts, they will lower the forecast. The two important changes to the manager's reporting strategy occur at the lower end of the manipulation range (region with red stripes) and at the higher end (region with green stripes). Changes in forecast error for the regions below the red striped one and above the green striped ones balance each other. Between the red and green region, the forecast error will be zero.

Thus, we can obtain the following formula for the change in squared error:

$$E[(z_o - y)^2] - E[(z_m - y)^2] = \int_{z_o}^{z_m} (x - z_o)^2 f(x)dx - \int_{z_o - m}^{z_m - m} (z_m - x)^2 f(x)dx. \tag{2.6}$$

The first summand is the increase in error at the high end of the manipulation range. Originally, there was no error because the manager would manipulate to the target. With the lowered forecast, he no longer has to manipulate when earnings fall into this range, and there will be a positive error. The second summand is the decrease in error at the low end of the manipulation range. With the lowered target, the manager will now manipulate to the target and the error is zero. Originally, earnings were too low at this point and the manager did not manipulate, leading to a positive error. One can see that the error in the first term will be close to zero because with a lower target, reported earnings are only a smidgen above the forecast. However, the error in the second term is positive because it is at least m. Formally, using a first-order approximation approach shows that the first term approaches zero, while the second term is nonzero. Thus, the squared error is decreasing at z_m, proving that $z_o < z_m$ is correct, which is summarized in the following lemma.

Lemma 2.1 The fewer analysts follow a firm, whose forecasts are used as a target for the manager, the lower the forecast will be.

Intuitively, the manipulation range is beneficial to the analysts, because it eliminates the forecast error compeletely. If actual earnings fall into this range, then the manager manipulates to the forecast, and the error is zero. The analyst, if there are not too many following the same firm, can influence the manager's reporting strategy and hence the location of the manipulation range. Naturally, he would like to eliminate bigger, rather than smaller, forecast errors. Essentially, the analyst lowers his forecast to allow the manager to manipulate to the target for lower earnings realizations than before.

It should be noted that if the analysts believe that the manager's target is the median forecast, then there will be no change in equilibrium, when there are three or more analysts. This is because no single analyst can change the median by changing his forecast, when there are at least three analysts. Although, if one were to introduce forecast dispersion into the model, then each analyst, ex-ante, would have a $\frac{1}{N}$ chance (N is number of analysts) of issuing the median forecast. Thus, each analyst would anticipate that a change in his forecast would matter in expectation by $\frac{1}{N}$ of that change, and results would be the same as when the manager's target is the mean forecast.

Now we can go back to equation (2.4) and evaluate how manipulation changes when the forecast decreases. The width of the manipulation range (m) and the amount of manipulation ($z - x$) are unchanged. However, the probability that manipulation occurs will be different. If the mean of the earnings distribution is on the downward sloping part (mean higher than mode), then a forecast decrease will increase the probability of manipulation. Vice versa, if the mean is on the upward sloping part of the distribution, then a forecast decrease will decrease the probability of manipulation. The mean of a distribution is higher than the mode, if it is right-skewed (or positively skewed) using the Pearson mode skewness. Formally, we can see the above intuition taking a derivate of equation (2.4) and applying the Leibniz rule:

$$\frac{d}{dz}\left(\int_{z-m}^{z}(z-x)f(x)dx\right) = F(z) - F(z-m) - mf(z-m). \tag{2.7}$$

The remainder of the proof again uses first-order approximation. Thus, we can state the above result in the following proposition.

Proposition 2.1 A lower analyst following leads to more expected manipulation if the earnings distribution is right-skewed and vice versa.

The above result shows that the number of analysts covering a firm can influence that firm's manager's earnings management decisions, even when there is no monitoring occuring. Prior archival studies (Yu 2008, Chen et al. 2015) on this topic have attributed the relationship to the monitoring hypothesis without accounting for potential effects of skewness. Gu and Wu (2003) find that the median skewness is about zero, but there is "large cross-sectional variation in earnings skewness." Hence, it seems plausible that average skewness in other studies could be positive or negative, depending on the specific sample that is used. Therefore, I propose that skewness should be a control variable in future research to clearly identify the forces driving the relationship.

In the model, analysts are completely rational. Yet, from an outsider's perspective, it can appear as if a lone analyst does not minimize the forecast error, if one mistakenly assumes the manager's reporting strategy is constant. Archival researchers can observe that expected value (or median for that matter) of reported earnings and analyst forecast do not match when there are few analysts (Brown 1997, Lim 2001, Hong and Kacperczyk 2010). However, this bias is not real and perfectly reasonable from the analyst's point of view. The analyst not only minimizes the error given the distribution, but also takes the indirect effect, by changing the manager's manipulation strategy, into account.

2.4.3 Comparative Statics

The model allows for more insights to be gained about the effect of analyst coverage on accounting manipulation.

Corollary 2.1 An increase in variance of the earnings distribution leads to

(i) less expected manipulation,

(ii) a decreased impact of a change in analyst following on manipulation.

The first one follows from the fact that an increase in variance makes it less likely that actual earnings fall just below the forecast target. Most of the time, earnings will either be high enough such that manipulation is not necessary, or low enough such that too much manipulation would be required (more than the manager is comfortable with). Intuition for (ii) is as follows: Since manipulation is less likely to occur, analysts will be less concerned about a change in their forecasts affecting the manager's reporting strategy. This effect causes forecasts to increase and approach the "many analysts equilibrium". Another effect happens, if analysts forecast the median, because a mean-preserving spread changes the median, as long as the distribution is skewed. For example, for a right-skewed distribution, an increase in variance leads to a decrease in the median (see Figure 2.3).

Corollary 2.2 When the earnings distribution becomes more skewed (deviates more from zero skewness), then a change in analyst following causes a larger change in earnings management.

The intuition here stems from the fact that there is no effect of a change

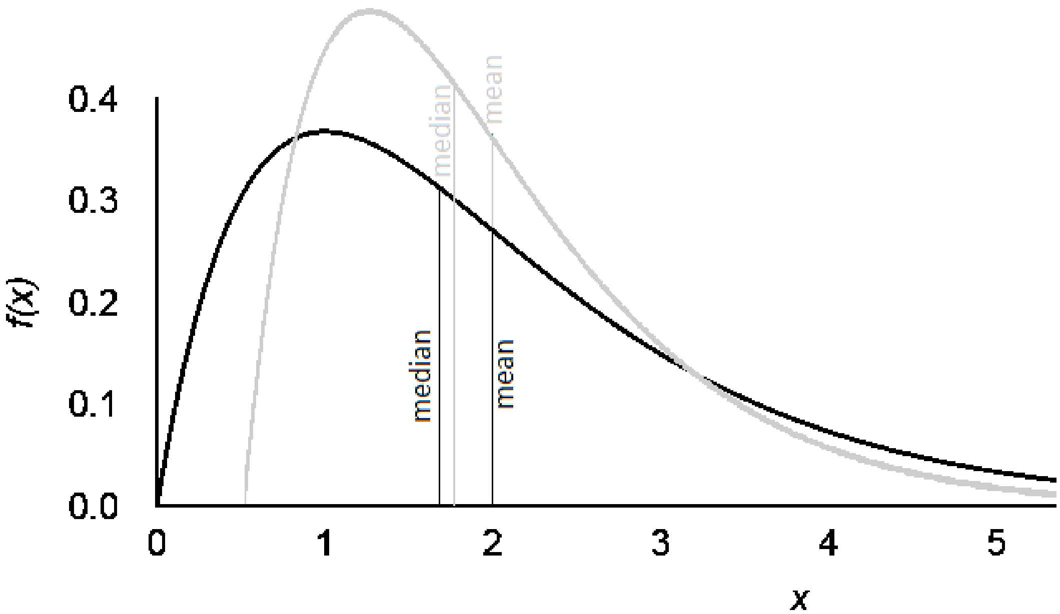

Figure 2.3: The black line is the same probability density function as in the previous figures: $f(x) \sim \Gamma(2, 1)$. the grey line is $f(x + 0.5) \sim \Gamma(2, 0.75)$. For both distributions, $E[x] = 2$. However, the median of the grey distribution with less variance (1.76) is higher than the median of the black distribution (1.68), illustrating the manager's incentives to decrease variance, when the distribution is right-skewed.

in analyst following on manipulation, if the distribution is not skewed at all. There will still be an effect on the average forecast, but a change in the forecast does not lead to a change in expected manipulation because the probability density function is flat around the mean (and median). As skewness deviates more from zero, this does not persist and the probability of manipulation changes with a change in the forecast (i.e. the first derivative of the probability density function is nonzero).

Corollary 2.3 An increase in the manipulation range m causes an increase in analysts' forecasts.

Analysts try to forecast reported earnings, not actual earnings. If manipulation is easier for the firm manager, then expected reported earnings increase, and so should the analysts' forecasts, i.e. $\frac{dz}{dm} > 0$. On the other hand, it is also instructive to analyze, whether forecasts will increase by more than the increase in m. Clearly, this cannot be optimal for the analysts. Manipulation is not guaranteed to occur, so reported earnings increase by less than the increase in m. Thus, combining these observations, it has to be true that $0 < \frac{dz}{dm} < 1$.

2.4.4 Discussion of Results

The above result shows that the number of analysts covering a firm can influence that firm's manager's earnings management decisions, even when there is no monitoring or pressure occuring. Prior archival studies on this topic have attributed the relationship to the monitoring hypothesis (Yu 2008, Chen et al. 2015) or the pressure hypothesis (He and Tian 2013, Huang et al. 2017) without accounting for potential effects of skewness. Gu and Wu (2003) find that the median skewness is about zero, but there is "large cross-sectional variation in earnings skewness." Hence, it seems plausible that average skewness in other studies could be positive or negative, depending on the specific sample that is used.

Gu and Wu (2003) also show that there is a positive correlation between firm size and skewness, which means that large firms are more likely to have a right-skewed distribution. This leads to the prediction that samples that are more heavily weighted towards large firms, would lead researchers to proclaim consistency with the monitoring hypothesis, while samples with a larger proportion of smaller firms would lead to researchers finding support for the pressure hypothesis. Yu (2008) only reports market value and Huang et al. (2017) only report the logarithm of market value. Log mean and mean log are not equivalent transformations. Thus, I am not able to compare the average firm size of these two papers. However, both Chen et al. (2015) and He and Tian (2013) report total assets. Indeed, the average firm size in the former paper is significantly bigger than in the latter,[2] consistent with my prediction. Therefore, I propose that skewness should be a control variable in future research to clearly identify the forces driving the relationship.

Additionally, Gu and Wu (2003) demonstrate a negative correlation between skewness and loss firms. This is logical because a loss is more likely to occur when the earnings distribution exhibits a long left tail. Hence, I would expect these firms to have a positive correlation between analyst following and earnings management (what archival researchers would label as consistent with the pressure hypothesis).

Finally, in the model, analysts are completely rational. Yet, from an outsider's perspective, it can appear as if a lone analyst does not minimize the forecast error, if one mistakenly assumes the manager's reporting strategy is constant. Archival researchers can observe that expected value (or median

[2]Chen et al. (2015) report mean total assets of \$10 billion and a median of \$2 billion. He and Tian (2013) report mean total assets of \$3.6 billion and a median of \$0.4 billion.

for that matter) of reported earnings and analyst forecast do not match when there are few analysts (Brown 1997, Lim 2001, Hong and Kacperczyk 2010). However, this bias is not real and perfectly reasonable from the analyst's point of view. The analyst not only minimizes the error given the distribution, but also takes the indirect effect, by changing the manager's manipulation strategy, into account.

2.5 Investment

In this section, I assume that earnings management is not possible, i.e. $m = 0$. On the other hand, the manager will now choose the investment level k endogenously. Thus, the complete focus of this section is on the effect of forecast beating on investment choice.

There are several relevant factors that affect the equilibrium investment level, neither of which is the number of analysts. In an equilibrium, neither the manager nor the analysts can have an incentive to deviate from their strategies. However, the key difference to the earnings management case above is that the manager makes the investment decision *before* the analysts decide on their forecast. Thus, regardless of the number of analysts, they will always take the investment level as given. This can be stated in the following observation.

Observation 2.1 The number of analysts does not affect the equilibrium investment level.

The three factors that are relevant include the analysts' loss function, earnings skewness, and investment observability. If analysts minimize the squared forecast error, then they will forecast the mean. If they minimize the absolute error, then they will forecast the median. Combined, these three factors produce the following proposition.

Proposition 2.2 If shape parameters are held constant, then the following statements are valid.

(i) If investment is observable, then investment will be first-best.

(ii) If investment is unobservable, and analysts minimize the squared error, then investment will be first-best.

(iii) If earnings are skewed, investment is unobservable, and analysts minimize the absolute error, then there will be underinvestment, if earnings are right-skewed, and overinvestment, if earnings are left-skewed.

The manager wants to maximize the probability of beating the forecast target. If earnings are not skewed, then he will beat the forecast exactly 50% of the time if he chooses first-best investment. Since he has no way of improving this probability, even off the equilibrium path, he has no incentive to deviate from first-best without earnings skewness. Part (i) follows a similar logic. Regardless of his choice, he will always beat the forecast a fixed percentage of the time. Due to observability, analysts adapt their forecasts such that the median/mean equals the forecast. Due to the assumption of constant shape parameters, the percentile of the mean is constant, and hence the manager's chances of outperforming the mean are as well. Counterintuitively, the manager does have an incentive to deviate from first-best, when investment is unobservable (part iii), when analysts forecast the median. In equilibrium, he will still beat the forecast target only 50% of the time. However, the allure of a higher success rate off the equilibrium path incentivizes him to deviate from first-best. At the first-best investment level, the mean of the earnings distribution will not change much when investment changes (the first derivative is zero). Yet, the median does change because variance changes. The difference between mean and median is proportional to the standard deviation of a distribution when shape parameters are held constant. The manager has an incentive to increase the median, because he takes the analysts' forecasts as given, who cannot adapt to changes in the manager's decision due to unobservability. When the distribution is left-skewed, then the median is higher than the mean. Hence, the manager wants to increase variance by overinvesting, which causes the median to increase and diverge from the mean. The reverse is true for a right-skewed distribution (part iii). Figure 2.3 shows the desribed behavior.

This effect is also known as the Red Queen effect (from Alice in Wonderland). The manager deviates from first-best, even though in equilibrium, he will still not have achieved anything. The analysts perfectly anticipate this deviation. The intuition for unobservable investment changes, when analysts forecast the mean (part ii). With constant shape parameters, the manager cannot influence the percentile ranking of the mean. The only way of increasing his probability of success would be to increase the mean itself. This, however, will cause him to make the first-best investment decision.

So far, I kept the shape parameters constant and only allowed the manager to affect location and scale parameters. Realisticallly, a CEO will also have control over skewness itself. He could decide to invest in a long-shot project that succeeds only a small percentage of times (but pays off big when it does) or opt for a more balanced approach with equal likelihood of deviations from the mean. Take the following example: let earnings be distributed in the form of a demeaned Gamma distribution, with an extra term such that first-best is achieved at $\alpha = \alpha_{FB} = 2$:

$$x \sim \Gamma(\alpha, 1) - \alpha - (2 - \alpha)^2. \tag{2.8}$$

a is the shape parameter of the Gamma distribution and skewness is $\frac{2}{\sqrt{\alpha}}$. If the manager can choose α, then it is easy to show that he will choose $\alpha < \alpha_{FB}$, if analysts minimize the absolute error and investment is unobservable. A lower α implies a more positive skewness. Hence, the manager, given this opportunity set, has a long-shot bias. The bias, however, is not caused by the direct effect on skewness, but via an indirect effect on variance. One cannot change skewness of a Gamma distribution without also affecting variance. An increase in skewness causes a decrease in variance. The above described intuition that the manager prefers less variance when the distribution is right-skewed still applies, and thus results in a preference for more skewness. This intuition can be stated in the following observation.

Observation 2.2 If analysts minimize the absolute error and investment is unobservable, then depending on the opportunity set of the manager, he may exhibit a long-shot bias.

The manager's desire to beat the analysts' forecast can rationally explain the long-shot bias. Schneider and Spalt (2016) find that "capital expenditure is increasing in the expected skewness of segment returns." They attribute this result to the fact that managers use gut feel, which causes them to subjectively assess probabilities, overweighting low probability events, as predicted by prospect theory (Kahnemann and Tversky 1979).

2.6 Manipulation and Investment

Finally, I turn to the full model (as described in the model section) with an endogenous investment level k and a positive manipulation range width $m > 0$. The above observed forces of the previous two sections are of course still at play in the full model. Hence, this section will focus on the effects that the interaction between investment and earnings management causes.

The model can be solved backwards. When the manager makes the manipulation decision, he takes his own investment decision and the analysts' forecasts as given. When the analysts decide on their forecasts, they take the manager's investment decision as given. Hence, this part of the model is equal to the analysis of section 4. However, when the manager makes his investment decision, he will take the subsequent effects on analyst forecasts and earnings management into account. As it turns out, the interaction of investment and manipulation creates an incentive for the manager to decrease investment. Less investment is associated (by assumption) with less variance. Less earnings dispersion makes manipulation more effective because income will fall into the manipulation range more often, as the width of the range is exogenously fixed. The manager will not minimize variance completely, though. If investment is unobservable, then there is still an incentive to keep variance above the minimum, because this will positively affect median (and mean) earnings, allowing him to beat the target more often (off the equilibrium path, analysts of course anticipate this behavior in equilibrium). With observable investment, he would indeed minimize variance. This can easily be fixed, if one assumes that the executive not just cares about beating the forecast (as is done in this paper for simplicity), but also about firm value. The above described intuition can be summarized in the following proposition.

Proposition 2.3 The interaction between earnings management and investment causes the CEO to choose less investment and leads to more expected earnings management.

While Proposition 2.3 discusses the interaction effect, we can also study the combined effects discussed in Proposition 2.2 and 2.3. For this exercise, I define investment efficiency (IE):

$$IE(k) = -|k - k_{FB}|. \tag{2.9}$$

This expression yields a maximized investment efficiency for first-best investment $k = k_{FB}$. It follows that $IE(k_{FB}) \equiv IE_{FB} = 0$.

For an exogenously assumed right-skewed distribution, the possibility of earnings management unambiguously causes investment efficiency to deteriorate: $\frac{dIE}{dm} < 0$. The effects of proposition 2.2 and 2.3 operate in the same direction. However, this is not the case for a left-skewed distribution. A small amount of manipulation improves investment efficiency. There even exists a manipulation width ($m = m_{FB}$), such that investment efficiency reaches first-best. This will occur, when the direct effect described in section 2.5 and the interaction effect in section 2.6 exactly offset each other. Only when the manipulation range width exceeds the first-best optimal level ($m > m_{FB}$) will a wider range result in lower investment efficiency. This leads to the following propositon.

Proposition 2.4 If the earnings distribution is right-skewed, then an increase in the width of the manipulation range leads to lower investment efficiency: $\frac{dIE}{dm} < 0$. If earnings are left-skewed, investment is unobservable, and analysts minimize the absolute error, then there exists a manipulation range width $m = m_{FB}$, such that $IE = IE_{FB}$, and an increase in the deviation from that width decreases efficiency, $\frac{dIE}{d|m-m_{FB}|} < 0$.

The CEO wants to maximize the chances of meeting or beating the forecast. With earnings management, the important cutoff for this endeavor is $z - m$, the lower edge of the manipulation range. If earnings are above this threshold, then he will succeed, either because earnings are high enough on their own, or because manipulation elevates reported earnings to z. As the manipulation range width m increases, $z - m$ will decrease (z increases at a slower rate than m, see appendix). At some point for a left-skewed distribution, $z - m$ will be equal to $E[x]$, expected unmanipulated earnings. If the manager cannot influence shape parameters, then at that point he cannot improve his chances by deviating from first-best, because by definition $E[x]$ is maximized at first-best. As the manipulation potential grows even larger ($m > m_{FB}$), the lower bound of the range will be below mean unmanipulated earnings. Figure 2.4 illustrates this.

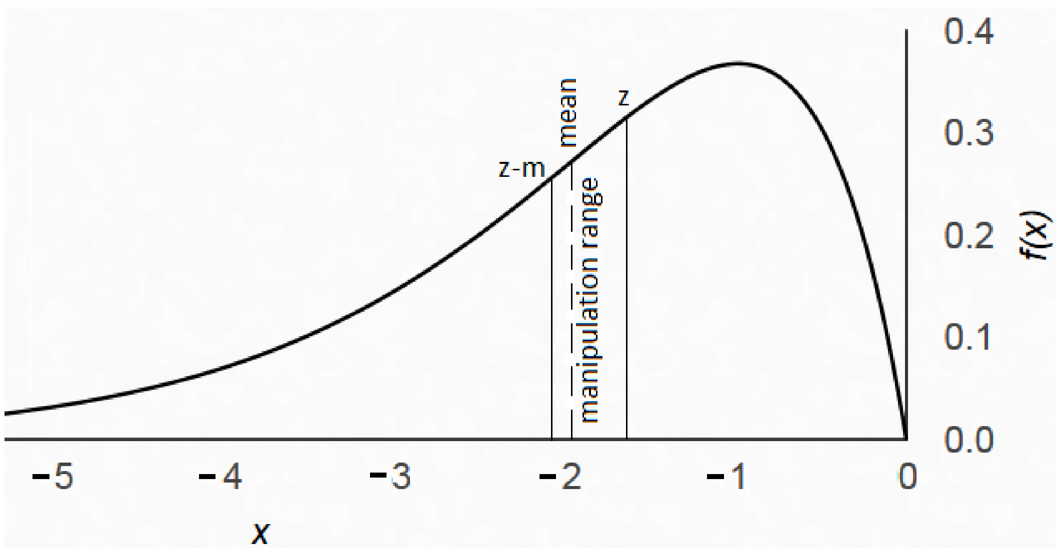

Figure 2.4: In this example, unmanipulated earnings x are distributed in the shape of a mirrored Gamma distribution: $f(-x) \sim \Gamma(2,1)$. It shows an example, for which the lower edge of the manipualtion range, $z - m$, is below the mean of the distribution. Thus, when m is large, even a left-skewed distribution can incentivize the CEO to reduce variance.

The managerial incentives then behave as if the distribution were right-skewed. A decrease in variance causes some probability mass to shift to the right of $z - m$, thus allowing the manager to manipulate those earnings realizations up to the forecast.

2.7 Conclusion

Analysts try to accurately predict reported earnings, which are influenced by managers' investment and reporting decisions. Executives, when making these decisions, try to increase their chances of beating the average analyst forecast. Thus, both sides influence (and are influenced by) the other's decision making process, which is the focus of this paper. I obtain several results. First, the number of analysts matters when it comes to the manager's manipulation decision. When there are many analysts, each one has only a small influence on the manager's target. Thus, they take the reported earnings distribution as given, and forecast the mean of said distribution. When there are few analysts, each anticipates that a change in their forecast affects the manager's behavior via a change in the target. Taking this into account, analysts lower their forecasts, compared to when there are more analysts. This change due to fewer analysts results in more expected earnings management when the distribution is right-skewed and less when it is left-skewed, providing an alternative explanation to both the monitoring hypothesis and the pressure theory postulated by archival studies, potentially explaining conflicting empirically observed results.

Second, the manager's investment decision can be affected. When analysts try to minimize the absolute forecast error and investment is unobservable, then the executive may deviate from first-best. The median of a distribution can improve, even when the mean remains virtually unchanged. The CEO can cause this, if investment affects the variance of the distribution. When the distribution is left-skewed, an increase in variance is beneficial to the agent, when it is right-skewed, he prefers a decrease. Depending on the opportunity set of the agent, this can rationally explain the empirically observed CEO long-shot bias.

Finally, I show that the interaction of investment and reporting decisions incentivizes the manager to underinvest (i.e. reduce variance) in order to improve the probability that earnings management allows him to beat the target. Thus, for left-skewed distributions, the possibility of earnings management can improve investment efficiency. The incentive to increase variance in order to improve the median, and the incentive to decrease variance to improve the efficiency of accounting manipulation can cancel each other out.

Combined, the findings of my paper provide a new perspective on several research questions. Future research can attempt to empiricially validate the predictions made in this paper.

Chapter 3

Accounting Conservatism and Convertible Debt

3.1 Introduction

One of the most important functions of accounting, maybe the most important, is providing external investors with financial information about the company that they plan to invest in or are already invested in. This creates an immediate and direct link between accounting and the investors' financing decisions. It should therefore come as no surprise that properties of accounting such as conservatism and precision influence this decision process. While these effects have been studied extensively for straight debt (e.g. Gigler et al. 2009, Gao 2013, Li 2013) and equity (e.g. Zhang 2000, Penman and Zhang 2002, Francis et al. 2004), convertible debt has been ignored. Lyandres and Zhdanov (2014) report that in a typical year, new issues exceed $100 billion globally, indicating that they are of economic significance.

The goal of this paper is to address two unresolved questions of the convertible bond literature: which theory best explains why firms issue convertibles and why are there differences in the announcement effect between countries. There are four main theories that provide an issuance motive: the risk shifting theory (Green 1984), the risk uncertainty theory (Brennan and Kraus 1987, Brennan and Schwartz 1988), the backdoor equity theory (Stein 1992), and the sequential financing theory (Mayers 1998). Yet, the archival and survey literature provides mixed results about the applicability of these four theories in the real world (see the literature review of Dutordoir et al. 2014). In this paper, I use the backdoor equity theory and am able to derive results that can answer the second question mentioned above, while the other three theories are not affected by accounting conservatism in the same way. I achieve this by assuming that the signal is not perfect and potentially upward or downward biased, thus departing from Stein (1992) and the perfect interim signal used there.

Specifically, I consider a model in which a firm needs to raise money from investors to fund a fixed-size investment opportunity. The firm knows whether the project is likely to succeed or not, but the market does not. If the company has a good project, then it wants to signal its type to the investors. It can accomplish this, if it picks a different financial instrument than bad firms. The available financing options are equity, convertible debt and straight debt. At an interim stage, the accounting system produces an observable signal, and subsequently investors can exercise their conversion option (to equity) in the case of convertible debt. If, at the very end, cash flows from the project are insufficient to repay debt, then the firm suffers a cost of financial distress.

In the benchmark case with a perfect signal, for a sufficiently high cost of financial distress, bad firms will issue equity while good firms choose convertibles. Investors will convert to equity after a high signal that reveals the firm to be good. This is optimal because then they will receive an uncapped share of the profits, outperforming debt. However, after a low signal, revealing a bad firm, investors would stick with debt, because the prospect of an uncapped profit share from high cash flows would be too dismal. This behavior guarantees that bad firms do not find it optimal to issue convertible bonds to mimic good firms as they would remain stuck on debt, subject to potential financial distress. Good firms achieve their ideal situation, because they are able to separate from bad firms, while not threatened by financial distress due to investors converting to equity.

These findings are not sustainable with an imperfect signal. The key problem is that in a perfectly separating equilibrium, investors would ignore the signal and go with their belief that only good firms issue convertible debt. As a result, they would always convert to equity. Unfortunately, this eliminates the threat for bad firms that they could get stuck on debt, and they now have no reason not to imitate the behavior of good firms, destroying the equilibrium. To solve this issue, some bad firms have to issue convertibles, making the signal relevant. In fact, the less imprecise the signal, the more bad firms have to switch from issuing equity to issuing convertibles. This is necessary because investors must find it optimal to not convert to equity after a bad signal. If the signal is less reliable, then the prior belief will be updated less, and a change in the behavior of bad firms affects the prior belief.

These changes have several implications. First, good firms bear all the costs of the decrease in reliability of the signal because, in equilibrium, bad firms will always be indifferent between their choices. Otherwise, a mixed strategy equilibrium would not be possible. Second, firms will experience financial distress more often. This holds for good firms due to the signal's increased likelihood of being false, and it holds for bad firms due to their changing behavior, there is no chance of distress after issuing equity. It should be noted that a certain accounting precision must be upheld. If the signal is too noisy, then one of two things will happen. Either the good firms do not find it optimal to issue convertibles and rather try to separate using straight debt, or the investors would not convert to equity, even after a high signal, thus eliminating any advantage that convertibles would have had.

I now turn towards the effects of conservatism. These will be the ones providing possible answers to the above mentioned questions. It is instructive to start with a very aggressive accounting system. In this case, it is possible to sustain a separating equilibrium. The important property of aggressive accounting is that a low signal is very informative. As a result, bad firms are not needed to ensure that investors do not convert to equity after a low signal. There is one important difference, though, compared to the case with a perfect signal. The incentive for bad firms to mimic good firms' behavior is increased due to the increased likelihood of a high signal, a more likely conversion to equity, and hence a lower distress probability. To counteract this, a higher cost of financial distress is necessary to keep incentives in balance. If this is satisfied, then the decreased informativeness of the high signal is irrelevant because the investors' prior belief makes them convert to equity after such a signal.

As the accounting becomes more conservative, the low signal becomes less informative. Thus, some similar forces as above with a decrease in accounting precision are at play. Investors need to stick with debt after a low signal. If, due to conservative accounting, many good firms end up reporting low earnings, then this can only be achieved if some bad firms issue convertibles. While these firms are kept indifferent, good firms have to bear the cost. These costs include more unfavorable financial terms because investors do not know that they are dealing with a good firm, and increased costs of financial distress driven by the increased chance of a low signal and investors not converting to equity.

These results can explain two unresolved questions of the convertible bond literature. When firms announce that they will issue convertible bonds, then there is an effect on the stock price. This effect is not uniform around the world. While the effect is significantly negative in the USA (Rahim et al. 2014), it is nonnegative or even positive in countries such as Japan (Kang and Stulz 1996), Taiwan (Chang et al. 2004), and the Netherlands (De Roon and

Veld 1998). These differences can be explained by differences in accounting conservatism. The USA has very conservative accounting in contrast to the other countries mentioned above (e.g. Weetman and Gray 1991, Ball et al. 2000, Li 2015). According to the results of my paper, investors will correctly assume that the average quality of issuers of convertibles is inferior when accounting is conservative, which will be reflected in the market capitalization. In my model, there can only be a positive announcement effect for convertibles. However, this can easily be fixed by going back to a model with three types (good, medium, and bad firms) as was the case in Stein (1992). The other unresolved question is about the validity of four issuer motive theories: the risk shifting theory, the risk uncertainty theory, the backdoor equity theory, and the sequential financing theory. The fact that the backdoor equity model used in this paper can explain these differences between countries, gives further support to this theory. The other three theories provide (non-)results with respect to conservatism that are at odds with the observed pattern. Thus, it appears that the backdoor equity theory is the most applicable one in the real world.

The remaining paper is organized as follows. Section 3.2 discusses related literature, section 3.3 describes the model, section 3.4 analyzes the scenario with a perfect signal, section 3.5 discusses the main results, section 3.6 explores short-term debt and other issuance motives theories, and section 3.7 concludes. All proofs are in appendix C.

3.2 Related Literature

The convertible bond literature has received quite a bit of attention over the past few decades. The two streams that are most relevant for this paper are about issuance motives and shareholder wealth effects. There are four often-cited theories on issuance motives. The oldest one of these is the risk shifting theory of Green (1984). The risk shifting problem states that firms will shift to more risky strategies after they took on more debt, even if this does not maximize total cash flows. This is optimal for shareholders because a leveraged structure turns stocks into options. If cash flows fall below the face value of debt, then a further drop does not impact shareholders. However, they get the full value of any increase that exceeds the face value of debt. Debtholders anticipate this issue and demand a higher nominal interest rate. Convertibles can alleviate this problem because it gives debtholders the option to convert to equity if cash flows turn out to be higher than expected. Thus, original shareholders no longer receive the full increase, and the upside of risk shifting is diminished.

The second one is the risk uncertainty theory by Brennan and Kraus (1987). This theory argues that a firm's manager may have superior knowledge about the risk of the company than the investors do. If the manager knows that the risk is lower than the market's perception, then debt will be underpriced (and vice versa) from the manager's point of view. This could keep him from issuing debt and thus forgoing profitable investments. A convertible bond introduces an option to convert into equity. This option value will be overpriced, if the manager knows that firm risk is actually lower than perceived. Combining these two observations yields the result that over- and undervaluation will cancel each other out and the convertible debt is fairly priced, eliminating the asymmetric information cost.

The third one, which is the one that this paper spends most of its focus on, is the backdoor equity theory of Stein (1992). Stein makes the case that there is asymmetric information about the fundamental firm value and that financial instruments can be used as signalling devices to reduce adverse selection costs. Firms have a preference towards equity because that avoids the cost of financial distress. Convertibles can help solve this problem by simultaneously signalling firm type and avoiding financial distress. This is achieved by structuring the convertible bond in a way such that the conversion decision is made after information asymmetry is resolved. Good firms will issue debt because it is assumed that there is no potential risk of financial distress, medium firms will issue convertibles that investors always convert to equity, and bad firms will issue stocks.

The last and most recent theory is the sequential financing theory (Mayers 1998). Mayers uses two forces: issue costs and overinvestment incentives. Whenever a firm raises funds, it incurs a substantial amount of issue costs, some of which are fixed, some variable. This would imply that long-term debt is preferable to short-term debt. On the other hand, if a manager has funds available, he will use them even if the project has a negative net present value. This would imply that short-term debt is preferable to long-term debt. Convertible bonds can combine the best of both worlds to solve this problem. If the future project turns out to be profitable, investors will convert to equity, leave the funds in the firm and issue costs are minimized. If the project should not be executed, then investors insist on the debt repayment, do not convert to equity, withdraw the funds, and hence minimize the overinvestment problem.

The other stream of the convertibles literature is about shareholder wealth effects, and more specifically relevant for this paper the announcement effect. If the choice of financing instrument is a signal about firm characteristics, then one would expect the market to react to the announcement of a new issuance. Consistent with the predictions of Myers and Majluf (1984) and Stein (1992), Eckbo et al. (2007) find that the announcement effect of convertibles (-1.8%) is between the effect for straight bonds (-0.2% and non-significant) and seasoned equity (-2.2%). However, the effect is not uniform across the globe. The meta-analysis of Rahim et al. (2014) provides a comprehensive overview and shows that the effect is significantly more negative in the USA than in other countries. For some countries such as Japan, Taiwan, and the Netherlands, the announcement effect is even nonnegative.

The third and last stream of literature relevant for this paper is about accounting conservatism. Basu (1997) defines conservatism as "resulting in earnings reflecting 'bad news' more quickly than 'good news'." The most important analytical paper to combine conservatism and corporate finance is Gigler et al. (2009), who showed that the prevailing wisdom that conservatism is good for debt contracting does not necessarily have to be true. Internationally, there are however differences in the degree of conservatism. Consistently, the USA and its Generally Accepted Accounting Principles (GAAP) rank among the most conservative systems (e.g. Weetman and Gray 1991, Ball et al. 2000, Li 2015).

3.3 Model

The model is a simplified and extended version of Stein (1992). There is a risk-neutral firm and many risk-neutral investors, the interest rate is zero. The firm has a project for which it needs to raise funds from investors.

Timing: There are three dates: $t \in \{0, 1, 2\}$. At $t = 0$, the firm has an investment opportunity available. The project quality is known to the firm, but not to the investors. The firm needs to raise the funds and can do so by issuing either ordinary shares, straight long-term debt or convertibles. Short-term debt is considered in a separate section. At $t = 1$, the accounting system generates a signal that is informative about the project quality. After the signal, if convertible bonds had been issued, the investors can either convert to equity or stick with debt. At $t = 2$, the project cash flow is realized and investors receive their payment according to the held financial instrument. If the face value of debt exceeds cash flows, then the firm suffers a cost of financial distress, and the game ends. Figure 3.1 depicts the timeline.

Investment project: At $t = 0$, the firm has a fixed-size investment opportunity available. In order to pursue it, it needs to raise and invest capital I from external sources. Cash flow X from the project arrives at $t = 2$ and can

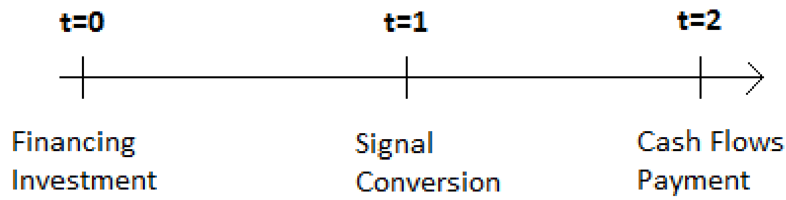

Figure 3.1: Timeline of the Model

be either high or low: $X \in \{X_H, X_L\}$. I use the notation $\Delta_X = X_H - X_L > 0$. The probability θ that the project generates high cash flows is known to the firm, but not to the investors and can either be θ_G or θ_B. For simplicity, I will refer to firms, whose project succeeds with the higher probability θ_G as good firms (G) and the other firms as bad ones (B). The investors prior belief about project quality is that a firm is good with probability v. Necessary restrictions to generate nontrivial solutions are $X_L + \theta_B \Delta_X > I > X_L$. The first inequality guarantees that bad firms can raise funds even when investors know that they are bad. The second inequality ensures that the face value of debt (which will be at least I) can exceed the cash flow.

Financial instruments: I restrict the analysis to the following three ways for the firm to raise funds at $t = 0$: ordinary shares, straight long-term debt, and convertible bonds. The case of short-term debt is addressed in a separate section. Ordinary shares represent the claim to a certain percentage Ψ_E of the cash flows, so investors receive $\Psi_E X_H$ or $\Psi_E X_L$ at $t = 2$. The second option is long-term debt. This debt can be described as a claim that pays the investors the lower value of the face value F_D and the cash flow X, i.e. $\min\{F_D, X\}$. Given the restriction $I > X_L$, the face value F_D will be greater than the low cash flow X_L and therefore creditors receive either X_L if cash flows are low or F_D if they are high. Furthermore, if the face value exceeds cash flows, then the firm experiences financial distress and suffers a cost C. This cost is nonnegotiable and personal. It can represent the time and money wasted on dealing with a looming bankruptcy. Lastly, the firm can issue convertible bonds. These are, if no conversion occurs, just like straight debt with a face value F_C and a claim to receive the lower value of F_C and cash flow X, i.e. $\min\{F_C, X\}$ at $t = 2$. However, at $t = 1$, the investors have the option to convert their claim into equity. Specifically, they can opt to receive a fraction Ψ_C of the firm's cash flows at $t = 2$, i.e. $\Psi_C X_H$ or $\Psi_C X_L$.

Accounting system: At $t = 1$, the accounting system generates a binary signal S that can either be high or low: $S \in \{S_H, S_L\}$ and provides information about the project's quality. The probabilities depend on project quality and are as follows:

$$
\begin{aligned}
p_{HG} &= \Pr(S_H|G) = 1 - (1-p)c, \\
p_{LG} &= \Pr(S_L|G) = (1-p)c, \\
p_{HB} &= \Pr(S_H|B) = (1-p)(1-c), \\
p_{LB} &= \Pr(S_L|B) = p + (1-p)c.
\end{aligned}
$$

This structure deviates slightly from prior literature (e.g. Venugopalan 2004, Chen et al. 2007, Bertomeu et al. 2016), but its interpretation is much the same. $p \in [0, 1]$ can be interpreted as accounting precision or informativeness. As p increases, both p_{HG} and p_{LB} increase and the signal is more likely correct. At the extremes, it is either perfectly informative (for $p = 1$) or completely uninformative (for $p = 0$). The deviation from prior literature is such that in this model a decrease in p affects both signals and not just the high one. $c \in [0, 1]$ is a measure of accounting conservatism or bias. As c increases, the probability of a bad signal increases for both fundamental types, and vice versa. Also, the probability structure satisfies the monotone likelihood ratio property, $p_{HG} \geq p_{HB}$.

3.4 Perfect Signal

In this section, I assume that the signal at $t = 1$ is perfectly informative about firm type, i.e. $p = 1$. The following is similar to the result obtained in Stein (1992) and is used as a benchmark to contrast the later results. With a perfect signal, a separating equilibrium can be achieved. Proposition 1 summarizes this equilibrium.

Proposition 3.1 For sufficiently high costs of financial distress (defined further below) and a perfectly informative signal ($p = 1$), the following is a sustainable separating equilibrium:

(i) bad firms issue a fraction $\Psi_E = \frac{I}{X_L + \theta_B \Delta_X}$ of equity and invest,

(ii) good firms issue convertible bonds with a conversion ratio $\Psi_C = \frac{I}{X_L + \theta_G \Delta_X}$, face value $F_C = \frac{1}{\theta_G}(I - (1 - \theta_G) X_L)$, and invest.

It is instructive in proving this equilibrium to analyze all three potential actors: good firms, bad firms, and investors. All of them have to find it optimal to behave in the way the equilibrium describes it.

Beginning with the investors' point of view: if they observe an equity issue, they know that it must be a bad firm. Therefore, they expect that this firm will generate a cash flow of $X_L + \theta_B \Delta_X$ on average. Logically, the issued shares will be worth $\Psi_E X$ or $\Psi_E(X_L + \theta_B \Delta_X)$ in expectation. Using the expression for Ψ_E from Proposition 1 shows that the shares are worth exactly I and the firm is able to finance its project. If investors observe a convertible bonds issue, then they know that the firm is a good one and the accounting signal will be high. At $t = 1$, investors have to make a decision. If they switch to equity, then their claim is worth $\Psi_C X$ or I. If they do not convert, then their claim is worth $\theta_G F_C + (1 - \theta_G) X_L$ which simplifies to I. Investors are indifferent and will switch to equity (alternatively, one could slightly reduce the face value to eliminate indifference) and the firm is able to finance its project. Even though, on the equilibrium path, the signal will never be low after a convertible issue, it is still necessary to analyze this path as this will be relevant for the perspective of a bad firm. In that case, investors would stick with debt as that claim would be worth more than the equity they would receive:

$$\frac{I}{X_L + \theta_G \Delta_X}(X_L + \theta_B \Delta_X) < \theta_B F_C + (1 - \theta_B) X_L. \tag{3.1}$$

The above inequality is intuitively true because investors are indifferent when they know the firm is good. If the firm is bad, then the probability of receiving the higher upside of equity decreases and the probability of the higher downside increases, making the left-hand side strictly lower than the right-hand side.

Good firms have no incentive to deviate because the equilibrium is the best case scenario for them. They are able to completely separate from bad firms while at the same time avoiding costs of financial distress. There is no way for them to improve upon this position.

Lastly, it has to be true that bad firms actually prefer issuing equity. If they follow the equilibrium strategy of issuing equity, then they will reap the full benefit of the project: $X_L + \theta_B \Delta_X - I$ because investors are held at break-even and there is no cost of financial distress. There is a benefit and a cost to deviating and issuing convertible bonds. The benefit is that investors would mistakenly assume that the firm is good and grant it better financing conditions, resulting in a lower face value of debt than they would otherwise be willing to give a bad firm. The cost is the potential cost of financial distress, if cash flows are insufficient to repay investors. This yields the following inequality condition that has to be satisfied:

$$\theta_B(X_H - F_C) - (1 - \theta_B)C \leq X_L + \theta_B \Delta_X - I, \tag{3.2}$$

which can be rearranged to show that costs of financial distress have to be sufficiently high:

$$(1 - \theta_B)C \geq \frac{1}{\theta_G} \left(\theta_G - \theta_B\right) \left(I - X_L\right) > 0. \tag{3.3}$$

The signal (e.g. accounting profits) is of utmost importance in the separating equilibrium. It gives investors some control and allows them to treat good and bad firms differently. Without a signal, bad firms could always mimic good firms' behavior and benefit from cheaper financing, subsidized by good firms. This exemplary illustrates that convertible bonds can only work when investors behave differently depending on the signal. If the signal is high, investors should find it optimal to convert to equity, and if the signal is low, they should optimally stick with debt. In the case of a perfect signal, this is not a problem as investors immediately and completely disregard any prior belief and update it with the information gained from the signal. It will become clear in the next section that this will not happen once the signal is no longer perfect.

3.5 Imperfect Signal

In this section, I no longer restrict the precision of the signal: $p \in [0, 1)$. As I alluded to at the end of the last section, it now no longer is the case that any prior belief will be disregarded once the signal arrives. This is best illustrated by the above described separating equilibrium. If an investor observes a convertible bonds issue at $t = 0$, then he correctly surmises that the firm is good. But now it is not impossible that the signal at $t = 1$ is low on the equilibrium path. Thus, he will deduce that the signal is incorrect, not change his belief about the firm's type, and convert to equity regardless. So far, this would not be a problem. However, the investor no longer changes his conversion decision based on the signal; the signal is meaningless. A bad firm would anticipate these consequences and rationally issue convertibles, benefit from cheaper financing conditions, and be guaranteed to not suffer financial distress. Clearly, the equilibrium is not stable and falls apart.

Since a completely separating equilibrium is not possible, some bad firms need to issue convertible debt, z (an unconditional probability). From the perspective of investors, two probabilities are important: the probability that a firm is good after a high signal,

$$\Pr(G|S_H) = \sigma_1 = \frac{(1-(1-p)c)v}{(1-(1-p)c)v + (1-p)(1-c)z}, \tag{3.4}$$

and the probability that a firm is good after a low signal,

$$\Pr(G|S_L) = \sigma_2 = \frac{(1-p)cv}{(1-p)cv + (p+(1-p)c)z}. \tag{3.5}$$

Given the fact that the signal is informative, it is the case that $\sigma_1 \geq \frac{v}{v+z} \geq \sigma_2$. A high signal improves the chance that the firm is good, while a low signal reduces that chance.

For an equilibrium to be stable, several conditions have to be met. First, investors have to find it optimal to convert to equity after a high signal:

$$\begin{aligned}\Psi_C(\sigma_1(X_L + \theta_G\Delta_X) + (1-\sigma_1)(X_L + \theta_B\Delta_X)) \geq \\ \geq \sigma_1(X_L + \theta_G(F_C - X_L)) + (1-\sigma_1)(X_L + \theta_B(F_C - X_L)). \end{aligned} \tag{3.6}$$

Second, investors must *not* convert to equity after a low signal:

$$\Psi_C(\sigma_2(X_L + \theta_G \Delta_X) + (1-\sigma_2)(X_L + \theta_B \Delta_X)) \leq$$
$$\leq \sigma_2(X_L + \theta_G(F_C - X_L)) + (1-\sigma_2)(X_L + \theta_B(F_C - X_L)). \quad (3.7)$$

Given the unambigious effect of an increase in the number of bad firms issuing convertibles (z) on σ_1 and σ_2, (3.6) and (3.7) create a lower bound, $\underline{z}$ and an upper bound $\overline{z}$. Hence, as I explain in the appendix, multiple equilibria are possible. Furthermore, from the investors' perspective, they need to break-even. If they would in expectation incur a loss, then they would not be willing to provide the funds for the investment project.

$$I(v+z) = v(1-p)c(X_L + \theta_G(F_C - X_L)) + z(p + (1-p)c)(X_L + \theta_B(F_C - X_L)) +$$
$$+ \Psi_C(v(1-(1-p)c)(X_L + \theta_G \Delta_X) + z(1-p)(1-c)(X_L + \theta_B \Delta_X)) \quad (3.8)$$

The left-hand side represents the expected funds for firms that will issue convertible debt. The first term on the right-hand side captures claims when investors stick with debt, while the second term captures claims when investors switch to equity.

I now turn to the firms' strategy. Starting with good firms, they need to find it optimal to issue convertible bonds. Three other options are available: issue long-term debt, issue equity, and not raising any funds. The first one, long-term debt, is tricky, because no firms on the equilibrium path that I focus on, use this financial instrument. The investors' beliefs can be defined freely. To simplify the analysis, I assume that investors believe a firm issuing long-term debt is a bad one. In this case, there is no incentive for anyone to go down this path. This assumption has to be revisited, if the model were to consist of three firm types and the best firms issue long-term debt. The other two options, though, are potentially better and must be ruled out. Not raising any funds, which gives the firm zero utility, will always be dominated by issuing equity, which gives the firm strictly positive utility. Thus, one inequality remains:

$$(1-(1-p)c)(1-\Psi_C)(X_L+\theta_G\Delta_X)+(1-p)c(\theta_G(X_H-F_C)-(1-\theta_G)C) \geq \\ \geq (1-\Psi_E)(X_L+\theta_G\Delta_X). \quad (3.9)$$

Lastly, there is the point of view of bad firms. As I showed above, for an equilibrium with imprecise signals to work, some bad firms have to issue convertibles in order to make the signal relevant for investors' decision making. This means that they have to be indifferent between issuing convertibles and equity. Otherwise, if they would strictly prefer one over the other, all of them would pick the same strategy.

$$X_L+\theta_B\Delta_X-I= \\ =(1-p)(1-c)(1-\Psi_C)(X_L+\theta_B\Delta_X)+(p+(1-p)c)(\theta_B(X_H-F_C)-(1-\theta_B)C) \\ (3.10)$$

The left-hand side is dramatically simplified because investors know that a firm issuing equity is bad. Thus, there is no mispricing and investors get their funds back, in expectation. Note also that the LHS is independent of the accounting system. This implies that regardless of accounting precision and conservatism, bad firms always receive the same utility. I summarize this insight in the following Lemma.

Lemma 3.1 In a mixed strategy equilibrium, the utility of bad firms is independent of the accounting system.

The above insight is crucial. Not only is bad firms' utility unchanged, investors are, by assumption, also always held at the same utility (i.e. break-even). Thus, any change in social welfare is completely borne by good firms. Since every firm always receives sufficient funds to invest and execute its project, the only change in social surplus is the cost of financial distress. The accounting system affects the overall cost of financial distress in two ways: directly via the probability of a low signal, and indirectly via the likelihood of a convertibles issue by bad firms. First, the direct effect: If conservatism increases, then the accounting is more likely to report low profits. After a low signal, investors stick with debt and do not switch to equity, thus giving rise to the financial distress problem, if cash flows turn out to be low. The indirect effect works as follows: When the accounting system becomes less precise and more conservative, then more good firms will report a low signal at $t = 1$. To keep investors from equity conversion, more bad firms have to issue convertibles. This ensures that a low signal is still likely enough to come from a bad firm. As more bad firms issue convertibles, and not equity, financial distress is increasingly probable. The above intuition is summarized in the following Proposition.

Proposition 3.2 As the accounting system becomes conservative ($c > 0$) or imprecise ($p < 1$), bad firms will issue convertible bonds ($z > 0$) as compared to the pure strategy equilibrium of Proposition 3.1. Also, the resulting societal increase in costs of financial distress is completely borne by good firms.

The insight from Proposition 3.2 can be used to more directly address the main questions of the convertible bonds literature that this paper tries to answer. An unanswered question is about the differential announcement effect of convertibles between countries. The announcement effect is the difference in firm value for the initial owners. At the beginning of the game, at $t = 0$, all firms are valued the same because no information about firm type is known to investors and current owners cannot credibly convey their superior knowledge. Thus, at $t = 0$ the initial shares are valued at

$$V_0 = X_L + (v\theta_G + (1 - v)\theta_B)\Delta_X - I. \tag{3.11}$$

The first factor of the second summand is the average probability of high cash flows. Costs of financial distress are not included due to the assumption that these are borne by the entrepreneur, and not by investors, who would buy some of the shares. The value of those initial shares after a convertibles issue rises to

$$V_C = X_L + (\frac{v}{v+z}\theta_G + \frac{z}{v+z}\theta_B)\Delta_X - I. \tag{3.12}$$

From (3.12) it becomes clear that an increase in the number of bad firms issuing convertibles (z) will decrease the firm value and hence the announcement effect (i.e. $V_C - V_0$). Together with the insight from Proposition 3.2, this shows that an increase in conservatism decreases the announcement effect. This is summarized in Propositon 3.3.

Proposition 3.3 When the accounting system is conservative ($c > 0$), then the announcement effect ($V_C - V_0$) is lower compared to the pure strategy equilibrium of Proposition 3.1.

This conclusion is critical to answering the question of why the announcement effect is different around the world. Dutordoir et al. (2014) write in their literature review: "[L]ittle is known about why convertible bond announcements made in countries like Japan and The Netherlands provoke non-negative stock reactions." And Rahim et al. (2014) in their meta analysis find: "Abnormal returns for hybrid securities issued in the United States are significantly more negative than those issued in other countries. [...] Finally, several factors identified as important by theory or in prior research are not significant within our cross-study models." In the literature, the USA is consistently among the countries with the most conservative accounting system (e.g. Weetman and Gray 1991, Ball et al. 2000, Li 2015), which explains why the announcement effect is more negative there. Similarly, the nonnegative and positive effects in Japan and the Netherlands can also be explained this way. As shown in the above mentioned prior literature, these countries have a very low amount of accounting conservatism. For example, Weetman and

Gray (1991) conclude: "[T]here is evidence to suggest that Dutch GAAP are at the less conservative end of the spectrum of financial reporting measurement practices." Naturally, a formal statistical test is required to confirm the observed pattern, which is beyond the scope of this paper.

3.6 Miscellaneous

3.6.1 Comparative Statics

There are basically three variables which combined determine the equilibrium. While the share of bad firms issuing convertibles (z) is certainly the most important for this paper, the other two, face value of debt, F_C, and conversion ratio, Ψ_C, shall not be neglected. The following observations can be made about their relationship.

Corollary 3.1 (i) If, and only if, the accounting is precise enough ($p > p^*$ as defined in the appendix), then the face value of the convertibles, F_C, is increasing as the number of bad firms issuing convertibles, z, is decreasing, i.e. $\frac{dF_C}{dz} < 0$.

(ii) If the face value of the convertibles, F_C, is increased, then the conversion ratio, Ψ_C, has to decrease, i.e. $\frac{d\Psi_C}{dF_C} < 0$, holding the accounting system constant.

The second insight is directly derived from the indifference condition for bad firms. They must always be kept indifferent between issuing convertibles and issuing equity. The utility derived from the latter is constant. when the conversion ratio increases, then these firms will benefit less as they get to keep less shares. To balance this out, the face value has to decrease, meaning that they will keep more funds after success. This result holds for a constant accounting system as a change in precision or conservatism of course would also affect the conversion probability.

This ties into the first observation. If bad firms are more likely to issue convertibles, then this adversely affects investors as they of course would prefer to lend money to a respectable firm. However, given corollary 1(ii), it is not possible to simply increase both the face value and the conversion ratio. To solve this problem, the financial terms have to become more beneficial to investors, while not hurting bad firms. Hence, good firms have to suffer. When the accounting is precise enough, then good firms are far more likely to achieve conversion after a high signal than bad firms. Thus, increasing the conversion ratio while decreasing the face value will hurt good firms while bad ones are indifferent. This pattern flips as the accounting becomes too noisy. The difference in conversion probability is small enough such that it is now no longer the dominant force. The pattern is dominated by the fact that bad firms are more likely to receive the low cash flow and they therefore care more about keeping more funds in that case, which will be the case when the conversion ratio is smaller, and hence the face value is higher.

3.6.2 Short-Term Debt

So far, short-term debt was exogenously excluded from the analysis. This, of course, is a little bit unsatisfying. In the model in its present form, short-term debt could reasonably fulfill the role played by convertible bonds. Firms, instead of using convertibles, issue short-term debt. At $t = 1$, they could then issue equity after they were able to distinguish themselves from bad firms via the accounting signal, assuming the signal was high. As Stein (1992) showed, a more complex model can avoid the separating equilibrium using short-term debt. If good firms receive additional information about its project's quality at $t = 1$, and if there is a liquidation option, then it is possible that firms that receive good news will choose not to raise new funds via an equity issue, because they are undervalued. Instead, they opt to liquidate some of their assets to repay the loan. However, this creates an ex-ante inefficiency due to the lack of commitment to an equity issue. Convertibles can contractually commit firms because it is investors, who decide whether to convert to equity or not, thus avoiding the ex-ante inefficient liquidation of assets. Stein shows that for a sufficiently low (but positive) net present value and a sufficiently high (but inefficient) liquidation value, short-term debt may not be optimal.

Another way of endogenously excluding short-term debt would be to introduce issue costs as in Mayers (1998). Assuming that raising funds involves variable and fixed costs, it is efficient to only raise funds once, instead of twice, which would occur with short-term debt. Costs for the services provided by investment banks, auditors, and stock exchanges can be nonnegligible. While short-term debt raised from a single bank is usually less costly, big debt obligations raised from bond markets have significant issue costs, too.

3.6.3 Three other convertibles theories

The model in this paper uses the backdoor equity theory of Stein (1992). There are, however, three other main issuance motives theories (see the related literature section). It is still an unresolved question in the literature which one of these best describes the real world. Dutordoir et al. (2014) conclude that empirical studies "do not reveal a clear pattern of evidence" for or against one of the four theories. As shown above, Stein's theory can explain and is consistent with an observed and previously unexplained pattern. Below, I will delve into the other three theories to determine what empirical predictions these generate with respect to conservatism.

First is the risk shifting theory of Green (1984). In that model, accounting does not play a role at all. The relevant problem does not occur due to asymmetric information, but due to a moral hazard. The firm, after securing debt, chooses a riskier, potentially slightly less profitable strategy at the expense of debt holders. Even if investors are fully informed, the moral hazard problem still persists, assuming that project selection or strategy choice is not contractible. Thus, it appears that the risk shifting theory does not explain the pattern of differential announcement effects across the globe.

Second is the risk uncertainty theory by Brennan and Kraus (1987). This theory claims that convertible bonds can be optimal when investors do not know the risk of the company. A convertible can be structured such that its debt component is overvalued (undervalued), while its equity component is undervalued (overvalued). An accounting signal, regardless of its downward or upward bias, could resolve some of the uncertainty about risk, but not all of it. Thus, convertibles still serve a role of basically threading the needle. Again, it seems unclear how conservatism would affect the likelihood of a firm using convertibles (other than affecting its features, e.g. the conversion ratio).

Lastly, there is the sequential financing theory of Mayers (1998). Part of this theory is the (exogenously assumed) moral hazard of overinvestment. However, accounting still plays a crucial role in this model because they can intervene in the interim by pulling their funds out of the firm. The question of whether conservatism affects the preference towards convertibles depends on whether one assumes that the manager has superior knowledge of the project's quality. The original paper did not need to specify this assumption as both investors and firm are perfectly informed once the decision to pull funds has to be made. If there is no asymmetric information with imperfect accounting, then convertibles always dominate. This changes if the firm has superior information. It is then possible that long-term debt is better than convertibles. The ranking depends on the severity of the overinvestment problem and the average project quality. If the manager has such a strong preference towards investment over no investment that he is willing to execute substantially negative net present value projects, then convertibles are better. The same is true if the project is rarely of high quality, again intensifying the overinvestment problem. The advantage of long-term debt is that the firm can use its superior knowledge and avoid inefficient fund withdrawal if they know that the project's quality is good. These two forces thus determine the effect of conservatism. If overinvestment is a concern and the average project quality is low, then convertibles are always better than long-term debt. However, if this is not the case, then a more aggressive accounting system benefits convertibles as this decreases the chances of an inefficient fund withdrawal. Empirical predictions from this model include that in countries and companies with better corporate governance (and thus a lower overinvestment problem) and at prosperous economic times, convertibles will be used more often, but it is not clear how this would affect the announcement effect as this is not a signalling model. The proof is in appendix C.

3.7 Conclusion

This study examines the effects of a signal generated by an accounting system on the use of convertible bonds. When the signal is perfect, then a separating equilibrium in which bad firms issue equity and good firms issue convertibles can be sustained because investors use the signal for their conversion decision. If the signal is high, they convert to equity to benefit from the increased upside; and when the signal is low, they stick with debt to better protect against the increased downside. However, once the accounting system becomes imprecise, a separating equilibrium breaks down. Investors would know that a firm issuing convertible debt is good, attribute a low signal to noise, and always convert to equity. Consequently, bad firms would anticipate this behavior and mimic the behavior of good firms. Thus, some bad firms need to issue convertibles in equilibrium when the signal is not perfect.

A similar logic applies to conservatism. An aggressive accounting system improves the quality of a low signal, thus reducing the necessity for bad firms to issue convertibles in order to keep investors from converting to equity. The reduced number of bad firms issuing convertibles simultaneously improves the perceived firm quality after a high signal. For conservative accounting, bad firms need to issue convertibles to offset the fact that investors would otherwise ignore a low signal and simply always convert to equity.

These results can help answer two unresolved questions of the convertible bonds literature. The first asks why the announcement effect of convertibles is different for different countries. The effect is significantly negative in the USA, larger than in other countries, consistent with the fact that US-GAAP prescribes a very conservative accounting. Additionally, the effect is nonnegative or positive in the Netherlands, Japan, and Taiwan, consistent with their more aggressive accounting. The second question is about the applicability of four main theories. This paper provides evidence that the backdoor equity theory by Stein (1992) can explain patterns that the other theories cannot. While the backdoor equity theory is consistent with the country-specific announcement effects pattern, the other theories are not.

Appendix A

Proofs for Chapter 1

Proof for Proposition 1.1

The agent's effort incentive constraint is

$(p_h+\gamma)qw_{SG}+(p_h-\gamma)(1-q)w_{SB}+(1-p_h-\gamma)qw_{FG}+(1-p_h+\gamma)(1-q)w_{FB}-c \geq (p_l+\gamma)qw_{SG}+(p_l-\gamma)(1-q)w_{SB}+(1-p_l-\gamma)qw_{FG}+(1-p_l+\gamma)(1-q)w_{FB}$

This can be rearranged to

$$q(w_{SG}-w_{FG})+(1-q)(w_{SB}-w_{FB}) \geq \frac{c}{p_h-p_l} \tag{A.1}$$

From this expression, it is clear that $w_{FG}=w_{FB}=0$ is optimal. As a result, the principal solves the following simplified problem:

$$\min_{w_{SB},w_{SG}} (p_h-\gamma)(1-q)w_{SB}+(p_h+\gamma)qw_{SG} \tag{A.2}$$

$$\text{s.t. } qw_{SG} + (1-q)w_{SB} = \frac{c}{p_h - p_l}. \tag{A.3}$$

Since both the objective function and the constraint are linear in the arguments, the solution will be a corner solution. Hence, I simply examine both possible solutions.

Case 1: $w_{SG} = 0$

$$w_{SB} = \frac{c}{(1-q)(p_h - p_l)}, \tag{A.4}$$

$$E[w] = \frac{(p_h - \gamma)c}{p_h - p_l}. \tag{A.5}$$

Case 2: $w_{SB} = 0$

$$w_{SG} = \frac{c}{q(p_h - p_l)}, \tag{A.6}$$

$$E[w] = \frac{(p_h + \gamma)c}{p_h - p_l}. \tag{A.7}$$

Comparing the two expressions for the expected wage shows

$$\frac{(p_h - \gamma)c}{(p_h - p_l)} < \frac{(p_h + \gamma)c}{(p_h - p_l)}, \tag{A.8}$$

because $\gamma > 0$ and therefore $w_{SB} > w_{SG}$ is optimal, proving Proposition 1.1.

I now quickly describe the proof for the risk-aversion case, which shows the trade-off that I mentioned in the text.

The principal's problem can be stated in the following Lagrange form.

$$\underset{w_{SB},w_{SG}}{Max}\ L = -(p_h-\gamma)(1-q)w_{SB} - (p_h+\gamma)qw_{SG} + \\ + \lambda(qu(w_{SG}) + (1-q)u(w_{SB}) - \frac{c}{p_h-p_l}) \quad \text{(A.9)}$$

Taking the derivatives with respect to the arguments yields

$$\frac{dL}{dw_{SB}} = -(p_h-\gamma)(1-q) + \lambda(1-q)u'(w_{SB}) = 0 \quad \text{(A.10)}$$

$$\text{and } \frac{dL}{dw_{SG}} = -(p_h+\gamma)q + \lambda qu'(w_{SG}) = 0. \quad \text{(A.11)}$$

After solving both for λ and using some algebra, I arrive at the following expression:

$$\frac{u'(w_{SG})}{u'(w_{SB})} = \frac{p_h+\gamma}{p_h-\gamma} > 1. \quad \text{(A.12)}$$

This fraction shows that $w_{SB} > w_{SG}$ given the standard features of a risk-averse utility function.

Proof of Proposition 1.2

The agent's utility function for high effort is

$$U = (p_h+\gamma)qw_{SG} + (p_h-\gamma)(1-q)w_{SB} + (1-p_h-\gamma)q(m_G w_{SG} + (1-m_G)w_{FG}) + \\ + (1-p_h+\gamma)(1-q)(m_B w_{SB} + (1-m_B)w_{FB}) - c - \\ - (1-p_h-\gamma)q\frac{gm_G^2}{2} - (1-p_h+\gamma)(1-q)\frac{gm_B^2}{2} \quad \text{(A.13)}$$

Since w_{FG} and w_{FB} will not incentivize effort for the same reasons as in the case above, they will be zero. This simplifies the above equation.

To get the manipulation decision, I take the first derivative of U with respect to m_G and m_B respectively:

$$\frac{dU}{dm_B} = w_{SB} - gm_B = 0, \tag{A.14}$$

$$\frac{d}{dm_G} = w_{SG} - gm_G = 0, \tag{A.15}$$

I also need an effort incentive constraint:

$$q((1-m_G)w_{SG} + \frac{gm_G^2}{2}) + (1-q)((1-m_B)w_{SB} + \frac{gm_B^2}{2}) \geq \frac{c}{(p_h - p_l)}. \tag{A.16}$$

Plugging (A.14) and (A.15) into (A.16), and solving for w_{SG} yields

$$w_{SG} = g(1 - \sqrt{1 - \frac{2}{gq}(\frac{c}{(p_h - p_l)} - (1-q)((1 - \frac{w_{SB}}{g})w_{SB} + \frac{w_{SB}^2}{2g}))}). \tag{A.17}$$

Plugging all of the above generated insights into (1.7) gives

$$\begin{aligned} E[m] = & \frac{w_{SB}}{g}(1 - p_h + \gamma)(1-q) + \\ & + (1 - \sqrt{1 - \frac{2}{gq}(\frac{c}{(p_h - p_l)} - (1-q)((1 - \frac{w_{SB}}{g})w_{SB} + \frac{w_{SB}^2}{2g}))})(1 - p_h - \gamma)q. \end{aligned} \tag{A.18}$$

Taking the derivative of $E[m]$ with respect to w_{SB}, and plugging (A.17) back in for simplicity results in

$$\frac{dE[m]}{dw_{SB}} = \frac{1}{g}(1 - p_h + \gamma)(1 - q) - \frac{1}{g}(1 - p_h - \gamma)\frac{g - w_{SB}}{g - w_{SG}}(1 - q). \qquad \text{(A.19)}$$

If $w_{SB} = w_{SG}$, then the equation simplifies to $\gamma > 0$. If $w_{SB} > w_{SG}$, then the derivate is even greater, hence proving the first part of Proposition 1.2.

To solve for the minimum, I set the equation equal to zero. To check, whether the solution is a corner solution, I set $w_{SB} = 0$, and solve for w_{SG}. I eventually arrive at the following condition for a corner solution

$$g > g_T = \frac{c(1 - p_h + \gamma)^2}{2\gamma q\,(p_h - p_l)\,(1 - p_h)}, \qquad \text{(A.20)}$$

proving the second part of Proposition 1.2.

Proof of Proposition 1.3

When $w_{SB} \neq w_{SG}$, then (A.14) and (A.15) show that when the benchmark is observable, the optimal manipulation choice is $m_B \neq m_G$. Hence, the restriction imposed on the manipulation decision by an unobservable benchmark that $m_B = m_G$ must decrease the agent's utility.

Proof of Proposition 1.4

The principal's problem can be stated in Lagrange form:

$$\underset{w_{SG},w_{SB},m_G,m_B}{Max} L = -((p_h+\gamma)qw_{SG}+(p_h-\gamma)(1-q)w_{SB}+(1-p_h-\gamma)qm_G w_{SG}+$$
$$+(1-p_h+\gamma)(1-q)m_B w_{SB}) + \lambda(q((1-m_G)w_{SG}+\frac{gm_G^2}{2})+$$
$$+(1-q)((1-m_B)w_{SB}+\frac{gm_B^2}{2})-\frac{c}{(p_h-p_l)})+\mu(w_{SB}-gm_B)+\nu(w_{SG}-gm_G). \tag{A.21}$$

Proof of part (i)

Assume that the solution will be $w_{SB} > 0$ and $w_{SG} = 0$.

Using this, (A.16), and (A.14), I get

$$2(1-m_B)gm_B + gm_B^2 = \frac{2c}{(1-q)(p_h-p_l)} \text{ and} \tag{A.22}$$

$$w_{SB} = (g - \sqrt{\frac{g}{(1-q)(p_h-p_l)}(g(1-q)(p_h-p_l)-2c)}). \tag{A.23}$$

These equations can be used to solve for the Lagrangian multipliers. Importantly,

$$\lambda = \frac{(p_h-\gamma)g + 2(1-p_h+\gamma)w_{SB}}{g-w_{SB}} \text{ and } \nu = 0. \tag{A.24}$$

Then, one has to check whether $\frac{dL}{dw_{SG}} < 0$ is indeed satisfied:

$\lambda + v < p_h + \gamma$ reduces after some algebra to

$$g > g^* \equiv \frac{2c}{(1-q)(p_h - p_l)(1-b^2)}, \tag{A.25}$$

$$\text{where } b = 1 - \frac{2\gamma}{2 - p_h + 3\gamma} < 1, \tag{A.26}$$

thus proving part (i) of Proposition 1.4.

Proof of parts (ii), (iii), and (iv).

Assume that the solution will be $w_{SB} = w_{SG} = w$.

Using simplified notation ($m_B = m_G = m$), I get the following results:

$$w = gm, \tag{A.27}$$

$$m = 1 - \sqrt{1 - \frac{2c}{g(p_h - p_l)}} \tag{A.28}$$

Again, one has to check whether $\frac{dL}{dw_{SG}} = \frac{dL}{dw_{SB}} = 0$ is indeed satisfied. The four Lagrangian derivatives are

$$\frac{dL}{dw_{SB}} = -\left(1-q\right)\left(m_B(1-p_h+\gamma) + p_h - \gamma\right) + \lambda\left(1-m_B\right)\left(1-q\right) + \mu = 0, \tag{A.29}$$

$$\frac{dL}{dw_{SG}} = -q\left(m_G(1-p_h-\gamma) + p_h + \gamma\right) + \lambda q\left(1-m_G\right) + v = 0, \tag{A.30}$$

$$\frac{dL}{dm_B} = -w_{SB}\left(1-q\right)\left(1-p_h+\gamma\right) - \lambda\left(w_{SB} - gm_B\right)\left(1-q\right) - g\mu = 0, \tag{A.31}$$

$$\frac{dL}{dm_G} = -qw_G\left(1-p_h-\gamma\right) - q\lambda\left(w_G - gm_G\right) - g\nu = 0. \tag{A.32}$$

Solving for the Lagrangian multipliers ν and μ reduces the system to two equations:

$$m_B(1-p_h+\gamma)+p_h-\gamma-\lambda\,(1-m_B) = (-w_{SB}\,(1-p_h+\gamma)-\lambda\,(w_{SB}-gm_B))\frac{1}{g}, \tag{A.33}$$

$$m_G(1-p_h-\gamma)+p_h+\gamma-\lambda\,(1-m_G) = (-w_G\,(1-p_h-\gamma)-\lambda\,(w_G-gm_G))\frac{1}{g}. \tag{A.34}$$

Plugging the above expressions for wage and manipulation in, solving for λ, reducing the system to one equation, yields the following results:

$$m = \frac{1}{2}, \tag{A.35}$$

$$g = \widehat{g} \equiv \frac{8c}{3(p_h - p_l)}, \tag{A.36}$$

proving part (iii) of Proposition 1.4.

Additionally, one can show that for $g > \widehat{g}$, it is the case that with benchmark-independent compensation, $\frac{dL}{dw_{SB}} > 0$, and $\frac{dL}{dw_{SG}} < 0$. And if $g < \widehat{g}$, then vice versa. This shows that for $g > \widehat{g}$, RPE is optimal, while for $g < \widehat{g}$, JPE is optimal, proving parts (ii) and (iv) of Proposition 1.4. The proof can also be obtained by invoking the intermediate value theorem, because the functions are continuous.

Proof of part (v):

For this proof, I will show that when $g \equiv \frac{2c}{p_h - p_l}$, the only (asymptotically) feasible compensation structure is IPE.

Assume that the (only feasible) solution is $w_{SG} = w_{SB} = g = \frac{2c}{p_h - p_l}$.

Plugging these values into equation (A.17) shows that this contract satisfies the incentive constraint with equality.

Plugging the above values into equations (A.14) and (A.15) yields $m_B = m_G = 1$. Due to the negative relationship between w_{SG} and w_{SB}, any deviation from IPE would result in either the invalid $m_B > 1$ or the invalid $m_G > 1$, proving part (v) of the proposition that IPE is the only feasible solution.

Proof of Corrolary 1.1

The use of RPE can by analyzed with the cutoff values $\widehat{g}$ and g^*. An increase in g^* means that RPE is used less because the range "max RPE" (part (v) in Proposition 4) gets smaller. An increase in $\widehat{g}$ means that RPE is also used less because the range "RPE" (part (iv) in Proposition 4) gets smaller.

Take the above derived formula for g^* :

$$g^* \equiv \frac{2c}{(1-q)(p_h - p_l)(1-b^2)}, \tag{A.37}$$

$$\text{where } b = 1 - \frac{2\gamma}{2 - p_h + 3\gamma} < 1. \tag{A.38}$$

It is immediately clear that an increase in b causes an increase in g^*. Thus, it is sufficient to take the derivative of b with respect to γ:

$$\frac{db}{d\gamma} = -\frac{2(2-p_h)}{(3\gamma - p_h + 2)^2} < 0. \tag{A.39}$$

$\widehat{g}$ is not affected by a change in γ, thus completing the proof of part (i).

We can do the same for p_h :

$$\frac{db}{dp_h} = -\frac{2\gamma}{(3\gamma - p_h + 2)^2} < 0. \tag{A.40}$$

There is another p_h term in the formula for g^*, but the effect of a change there has the same direction as the effect via b (that is, increasing the denominator), and hence the inverse relationship holds. $\widehat{g}$ and p_h are also inversely related, thus completing the proof of part (ii).

The third part of the corollary can be shown directly:

$$\frac{dg^*}{dc} = \frac{2}{(1-q)(p_h - p_l)(1-b^2)} > 0, \tag{A.41}$$

An increase in c also causes $\widehat{g}$ to increase, thus the effect is again unambiguous, completing the proof of part (iii) of Corrolary 1.1.

Proof of Corrolary 1.2

Plugging (A.14) and (A.15) into the IC constraint (A.16) gives the following constraint:

$$q((1-m_G)gm_G + \frac{gm_G^2}{2}) + (1-q)((1-m_B)gm_B + \frac{gm_B^2}{2}) = \frac{c}{(p_h - p_l)}. \tag{A.42}$$

Since this constraint must hold with equality, there must be an inverse relationship between m and g.

To show (ii), assume that the agent will, irrationally, ex-post, decide not to manipulate. Proving that expected wage declines in this case, is sufficient proof, because manipulation is declining in g as well. Thus, I want to show that

$$E[w] = (p_h - \gamma)(1-q)w_{SB} + (p_h + \gamma)qw_{SG} \tag{A.43}$$

is declining in g. Solving (A.17) for w_{SB} and plugging it into (A.43) expresses expected wage as a function of w_{SG} alone.

$$E[w] = (p_h - \gamma)(1-q)g\left(1 - \sqrt{\frac{1}{1-q}\left(1 - \frac{2c}{g(p_h - p_l)} - q\left(1 - \frac{w_{SG}}{g}\right)^2\right)}\right) + (p_h + \gamma)qw_{SG} \tag{A.44}$$

Since $\frac{dw_{SG}}{dg} < 0$, it is sufficient to show that $\frac{dE[w]}{dw_{SG}} < 0$ to prove part (ii) of the corrolary.

The proof of (iii) is explained in the text below the proposition. The agent could keep her manipulation decision constant and simply enjoy a higher rent. Any change in manipulation level must be to the benefit of the agent.

Proof of Corollary 1.3

Taking the derivative of (A.37) with respect to q yields

$$\frac{dg^*}{dq} = \frac{2c}{(p_h - p_l)(1-b^2)(1-q)^2} > 0, \tag{A.45}$$

proving the corollary ($\widehat{g}$ is unaffected by a change in q).

Proof of Propostion 1.5

Let the renegotiated contract be $w_n = \{w_{SGn}, w_{SBn}, w_{FGn}, w_{FBn}\}$. If the principal has all the bargaining power, then he will offer a new contract that makes the agent indifferent, if her project failed, and reject the new offer, if she succeeded.

$$w_{SGn} = w_{FGn} = m_G w_{SG} + (1 - m_G) w_{FG} - \frac{g m_G^2}{2} = \frac{w_{SG}^2}{g} - \frac{w_{SG}^2}{2g} = \frac{w_{SG}^2}{2g}, \quad \text{(A.46)}$$

$$w_{SBn} = w_{FBn} = m_B w_{SB} + (1 - m_B) w_{FB} - \frac{g m_B^2}{2} = \frac{w_{SB}^2}{g} - \frac{w_{SB}^2}{2g} = \frac{w_{SB}^2}{2g}. \quad \text{(A.47)}$$

Thus, the new contract does not affect the effort incentive constraint. The remainder of the proof proceeds similar to the proof of Proposition 4.

$$\underset{w_{SG}, w_{SB}}{Max}\, L = -((p_h + \gamma) q w_{SG} + (p_h - \gamma)(1 - q) w_{SB} + (1 - p_h - \gamma) q \frac{w_{SG}^2}{2g} +$$
$$+ (1 - p_h + \gamma)(1 - q) \frac{w_{SB}^2}{2g}) + \lambda (q w_{SG} (1 - \frac{w_{SG}}{2g}) + (1 - q) w_{SB} (1 - \frac{w_{SB}}{2g}) - \frac{c}{(p_h - p_l)}) \quad \text{(A.48)}$$

Assume that the optimal solution is $w_{SG} = 0$ and $w_{SB} > 0$.

$$\frac{dL}{dw_{SG}} = -p_h - \gamma + \lambda < 0, \quad \text{(A.49)}$$

$$\frac{dL}{dw_{SB}} = \frac{1}{g} (1 - q) (\lambda (g - w_B) - (w_B - g\gamma + g p_h + \gamma w_B - w_B p_h)) = 0. \quad \text{(A.50)}$$

Using the agent's incentive constraint yields

$$w_{SB} = g\left(1 - \sqrt{1 - \frac{2c}{g(1-q)(p_h - p_l)}}\right). \tag{A.51}$$

I need to check whether $\frac{dL}{dw_{SG}} < 0$ is indeed satisfied:

$$-\frac{1}{g}q\left(w_G + g\gamma + gp_h - \gamma w_G - w_G p_h\right) + \frac{q\lambda}{g}\left(g - w_G\right) < 0, \tag{A.52}$$

which, after plugging in (A.51) and some algebra, simplifies to

$$g > \frac{2c}{(1-q)(p_h - p_l)(1 - \frac{(1-2\gamma)}{(1+2\gamma)}^2)} = g_n^*, \tag{A.53}$$

Comparing g^* with g_n^* shows that $g^* > g_n^*$, proving the main part of Proposition 5. The last sentence of the proposition can be shown by assuming that wage payments are equal and examining the Lagrangian derivatives, which yield that $\frac{dL}{dw_{SB}} > 0$, and $\frac{dL}{dw_{SG}} < 0$ for all values of g when a solution exists, which shows that JPE is never optimal.

Appendix B

Proofs for Chapter 2

Proof of Lemma 2.1

When there are many analysts, they forecast z_m while taking the manipulation range as given: $[z' - m, z']$

$$E[(z_m - y)^2] = \int_{-\infty}^{z'-m} (z_m - x)^2 f(x)dx + \int_{z'}^{+\infty} (z_m - x)^2 f(x)dx. \tag{B.1}$$

Minimzing the above equation will occur when the derivative is zero:

$$\frac{d}{dz_m} E[(z_m - y)^2] = \int_{-\infty}^{z'-m} 2(z_m - x)f(x)dx + \int_{z'}^{+\infty} 2(z_m - x)f(x)dx = 0. \tag{B.2}$$

When there is one analyst, he forecasts z and the manipulation range changes when the forecast changes: $[z - m, z]$.

$$E[(z - y)^2] = \int_{-\infty}^{z-m} (z - x)^2 f(x)dx + \int_{z}^{+\infty} (z - x)^2 f(x)dx. \tag{B.3}$$

Again, minimizing this equation occurs when the derivative is zero. I evaluate the derivate at $z = z_m$.

$$\frac{d}{dz}E[(z-y)^2]|_{z=z_m} = \int_{-\infty}^{z-m} 2(z_m - x)f(x)dx + \\ + \int_{z}^{+\infty} 2(z_m - x)f(x)dx + m^2 f(z_m - m) = m^2 f(z_m - m) > 0. \quad \text{(B.4)}$$

The first and second summand combined are equal to the derivative of the forecast error when there are many analysts, and thus sum to zero, when evaluated at $z = z_m$.

Proof of Proposition 2.1 (completed)

The missing part of the proof is the first-order approximation.

$$F(z) - F(z-m) - mf(z-m) \approx = m[f(z - \frac{m}{2}) - f(z-m)], \qquad \text{(B.5)}$$

which is greater than zero, if $f'(z) > 0$. The first derivative of the probability density function at the mean (or median) is positive, when the Pearson mode skewness is negative.

Proof of Corollary 2.1

A mean-preserving spread means that $f(x = E[x])$ is decreasing. The probability density function becomes flatter, thus values close to the mean become less likely. Note that I assumed that m is small relative to the earnings distribution, which implies that the forecast z is close to the mean. Taking the derivative of (2.4) with respect to $f(x)$, evaluated at $x = E[x]$, proves the first part of the corollary.

The second part of the corollary immediately follows from the first part.

Proof of Corollary 2.2

When the distribution is unskewed, then the first derivative of the distribution at the mean is zero: $f'(x = E[x]) = 0$. Using this to analyze (B.5) shows that there is no effect of a change in the forecast on earnings management, when the distribution is unskewed. As skewness increases, $abs[f(z - \frac{m}{2}) - f(z - m)]$ increases as well, proving the corollary.

Proof of Corollary 2.3

Taking the derivative of (2.4) with respect to m immediately proves the corollary.

Proof of Proposition 2(iii)

In an equilibrium, the manager cannot have an incentive to deviate. This will be the case when he cannot improve upon the 50% success rate he will have because analysts forecast the median. Due to the assumption of constant shape parameters, it is the case that the nonparametric skew is constant: $S = \frac{\mu - \nu}{\sigma}$, where μ is the mean, ν is the median, and σ is the standard deviation. Thus,

$$\nu = \mu - S\sigma. \tag{B.6}$$

The manager tries to maximize the median. Thus, I solve for the first derivative with respect to k and examine it at $k = k_{FB}$.

$$\frac{d\nu}{dk}|_{k=k_{FB}} = (\frac{d\mu}{dk} - S\frac{d\sigma}{dk})|_{k=k_{FB}} = -S\frac{d\sigma}{dk}|_{k=k_{FB}}. \tag{B.7}$$

At the first-best investment level, the first derivative of the mean is zero. The first derivative of the standard deviation is positive for all k. Thus, if skewness is positive, then an increase in variance must lead to a decrease in the median, and vice versa for negative skewness.

$$sign(\frac{d\nu}{dk}|_{k=k_{FB}}) = sign(-S) \tag{B.8}$$

Proof of Proposition 2.3

The manager wants to maximize the probability that he surpasses the target. Earnings management increases this probability by:

$$\int_{z-m}^{E[x]} f(x)dx. \tag{B.9}$$

Note that for values above the mean, earnings management does not increase the success rate because without EM, the forecast would have been equal to the mean (the proof works similarly if analysts minimize the absolute error). More probability mass will be centered around the mean, when the variance of the distribution is smaller. Thus, less investment increases equation (B.9).

Proof of Proposition 2.4

The proof for a right-skewed distribution follows immediately from the results obtained in Propositions 2.2 and 2.3. The proof for a left-skewed distribution needs to show that for a large enough m, there will be underinvestment.

The manager wants to maximize his chances of meeting or beating the forecast:

$$\max \int_{z-m}^{\infty} f(x)dx. \tag{B.10}$$

If investment is unobservable, then he will take the forecast z as given (and thus $z - m$ as well). There exists $m = m_{FB}$, such that $z - m_{FB} = E[x]$. Intuitively, as m grows large, the analysts would not have a forecast error for earnings that are above the mean (if they keep $z - m$ above the mean), but would still have one for realizations below the mean. Thus, at some point it will be beneficial to the analysts, if the manipulation range extends below the mean, because there is no benefit for them to increase z even further: $\lim_{x\to\infty} f(x) = 0$.

The manager wants to decrease the percentile at $z - m$ (he wants to maximize 1 minus the percentile at $z - m$). For constant shape parameters, the percentile of the mean is constant. Thus, when $z - m = E[x]$, then the manager cannot increase his chances of success by deviating from first-best. When $z - m < E[x]$ (when $m > m_{FB}$), then a decrease in variance (i.e. underinvestment) causes the percentiles below the mean to approach the mean. Hence, for a fixed value below the mean such as $z - m$, the percentile will decrease.

Appendix C

Proofs for Chapter 3

Proof of Proposition 3.2 (continued) and Comparative Statics

Some of the proof is in the main text above. This is a continuation and supplement. An equilibrium must satisfy five equations: three investor-related, (3.6), (3.7), (3.8), one good-firms-related, (3.9), and one bad-firms-related, (3.10). One omitted equation is that the investors will break-even for an equity issue, too. Since still only bad firms issue equity, Ψ_E is unchanged from Proposition 1. Thus, three unknowns remain: Ψ_C, F_C, and z. Since there are only two equalities, but three inequalities, multiple equilibria are possible. Recall that investors and bad firms are indifferent between these, and good firms bear all the changes in profits. Thus, to simplify the analysis, I will focus on the best possible equilibrium, when (3.7) is satisfied with equality (i.e. $z = \underline{z}$). If this equilibrium is not stable, then no mixed-strategy one is, because (3.9) is easiest to satisfy in the equilibrium that is best for good firms. This leaves three equalities and three unknowns, hence a solution must exist.

Solving (3.10) for Ψ_C and plugging the solution into (3.8) yields an equation for F_C in terms of z:

$$
\begin{aligned}
F_C = (I(v+z) + (vp_{LG}\theta_G - vp_{LG} - zp_{LB} + zp_{LB}\theta_B)X_L - vp_{HG}G_E - zp_{HB}B_E + \\
+ \frac{vp_{HG}G_E + zp_{HB}B_E}{p_{HB}} - I\frac{vp_{HG}G_E + zp_{HB}B_E}{p_{HB}B_E} + \frac{vp_{HG}G_E + zp_{HB}B_E}{p_{HB}B_E} \\
(1-\theta_B)C - \frac{vp_{HG}G_E + zp_{HB}B_E}{p_{HB}B_E}p_{LB}\theta_B X_H) \\
(vp_{LG}\theta_G + zp_{LB}\theta_B - \frac{vp_{HG}G_E + zp_{HB}B_E}{p_{HB}B_E}p_{LB}\theta_B)^{-1}.
\end{aligned}
\tag{C.1}
$$

Taking the derivative of the above equation with respect to z yields

$$
\frac{dF_C}{dz} = \frac{B_E p_{BH}\left(X_L p_{BL} - B_E - C + B_E p_{BH} + C\theta_B + \theta_B X_H p_{BL} - \theta_B X_L p_{BL}\right)}{v\theta_B G_E p_{BL} p_{GH} - v\theta_G B_E p_{BH} p_{GL}}.
\tag{C.2}
$$

This derivative is positive iff $\theta_G B_E p_{BH} p_{GL} > \theta_B G_E p_{BL} p_{GH}$, which can be rearranged to show that it is satisfied if p is above a certain threshold, proving Corollary 3.1(i). The left-hand side is decreasing in p (approaching 0 as p approaches 1) and the right-hand side is increasing in p. Hence, a threshold p^* must exist such that both sides are equal.

The proof of Corollary 3.1(ii) follows immediately from (3.10), rearranging in terms of Ψ_C and taking the derivative with respect to F_C.

Returning to the main proof, (C.2) can be plugged into (3.7) to derive the following implicit function:

$$G(z) = (1 - (B_E - I - p_{LB}(\theta_B(X_H - F_C) - (1-\theta_B)C))\frac{1}{p_{HB}B_E})$$
$$(\sigma_2 G_E + (1-\sigma_2)B_E) - \sigma_2(X_L + \theta_G(F_C - X_L)) - (1-\sigma_2)(X_L + \theta_B(F_C - X_L)) = 0. \tag{C.3}$$

Thus, the implicit function theorem can be applied to this equation: $\frac{dz}{dc} = -\frac{dU}{dc}(\frac{dU}{dz})^{-1}$.

The proof that $z > 0$ for $p < 1$ and $c > 0$ is done by contradiction. Suppose that $z = 0$. Then $\sigma_1 = \sigma_2 = 1$, i.e. investors know that a firm issuing convertibles is good and the signal will not change this belief. Thus, investors will not use the signal and either always convert to equity or always stick with debt. If the always convert, then bad firms could mirror good firms' strategy with no punishment. If they never convert, then convertibles serve no purpose and could be substituted with long-term debt.

It remains to show that for $p < 1$ and $c = 0$, a perfectly separating equilibrium can be sustained. In this case $\sigma_1 > \sigma_2 = 0$. This means that after a convertibles issue and a subsequent off-equilibrium bad signal, investors would know that the firm is bad and not convert to equity. Good firms will never produce a bad signal and always achieve conversion, i.e. they have no incentive to deviate. The last thing to check is that bad firms will not deviate.

This condition can be expressed as follows:

$$X_L + \theta_B \Delta_X - I \geq (1-p)(1-\Psi_C)(X_L + \theta_B \Delta_X) + p(\theta_B(X_H - F_C) - (1-\theta_B)C). \tag{C.4}$$

As long as C is sufficiently high enough, this condition will be met. A high C does not affect good firms, who never suffer from financial distress. This shows that for $c = 0$, we can achieve a perfectly separating equilibrium, i.e. $z = 0$, and that for a conservative accounting system ($c > 0$), the number of bad firms issuing convertibles, z, is higher. Proposition 3.3 uses this insight, along with the expressions (3.11) and (3.12).

Proof of sequential financing theory extension

The manager has three investment opportunities and no funds as described below.

Investment 1 is available at $t = 0$, needs investment K_1, and has an NPV of $N_1 > 0$.

Investment 2 is available at $t = 1$, needs investment K_2, and has an NPV of $N_2 \in \{x, -x\}$, which will be x with probability q.

Investment 3 is available at $t = 1$, needs investment K_2, and has an NPV of $-x < -N_3 < 0$.

Issue cost for financing are and the manager is (exogenously assumed) prone to overinvestment. To fund the investments, he can either issue short-term debt, long-term debt, or convertible bonds. The manager learns the project quality with certainty at $t = 1$, while the investors receive a signal with the same structure as in the main model. Furthermore, suppose that the signal is meaningful, i.e. after a low signal investors prefer no investment, after a high signal investors do.

Profit from raising long-term debt: $N_1 + qx - (1-q)N_3 - C$.

Profit from raising short-term debt: $N_1 - C + q(1-(1-p)c)(x-C) - (1-q)(1-p)(1-c)N_3$

Profit from raising convertibles: $N_1 - C + q(1-(1-p)c)x - (1-q)(1-p)(1-c)N_3$

First note that convertibles always dominate short-term debt in this model. So we just need to compare the remaining two options.

After some algebra, onvertibles dominate long-term debt if

$$qx(1-p)c < (1-(1-p)(1-c))(1-q)N_3, \tag{C.5}$$

which can be rearranged to

$$c(qx-(1-q)N_3)(1-p) < p(1-q)N_3. \tag{C.6}$$

If $(qx-(1-q)N_3)$ is negative, then convertibles dominate regardless of the accounting system. If this factor is positive, then convertibles dominate if conservatism is below a certain threshold.

Bibliography

Aggarwal, R.K. and Samwick, A.A., 1999. Executive compensation, strategic competition, and relative performance evaluation: Theory and evidence. The Journal of Finance 54(6): 1999–2043.

Armstrong, C., Chau, J., Ittner, C.D. and Xiao, J.J., 2017. Internal Versus External Earnings per Share Goals and CEO Incentives. Working Paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2919478.

Bagnoli, M. and Watts, S.G., 2000. The effect of relative performance evaluation on earnings management: a game-theoretic approach. Journal of Accounting and Public Policy 19(4-5): 377-397.

Balakrishnan, R., Lin, H., and Sivaramakrishnan, S., 2016. On the Relative Efficacies of Ranking and Absolute Performance Evaluation Systems. Working Paper. Available at papers.ssrn.com/ sol3/papers.cfm?abstract_id= 2828956.

Ball, R., Kothari, S.P. and Robin, A., 2000. The effect of international institutional factors on properties of accounting earnings. Journal of accounting and economics 29(1): 1-51.

Banker, R.D. and Datar, S.M., 1989. Sensitivity, precision, and linear aggregation of signals for performance evaluation. Journal of Accounting Research 27(1): 21-39.

Bannister, J.W. and Newman, H.A., 2003. Analysis of corporate disclosures on relative performance evaluation. Accounting Horizons 17(3): 235-246.

Bartov, E., Givoly, D. and Hayn, C., 2002. The rewards to meeting or beating earnings expectations. Journal of accounting and economics 33(2): 173-204.

Basu, S., 1997. The conservatism principle and the asymmetric timeliness of earnings 1. Journal of accounting and economics 24(1): 3-37.

Basu, S. and Markov, S., 2004. Loss function assumptions in rational expectations tests on financial analysts' earnings forecasts. Journal of Accounting and Economics 38: 171-203.

Bertomeu, J., Darrough, M., and Xue, W., 2016. Optimal conservatism with earnings manipulation. Contemporary Accounting Research 34(1): 252-284.

Bertrand, M. and Mullainathan, S., 2001. Are CEOs rewarded for luck? The ones without principals are. The Quarterly Journal of Economics 116(3): 901-932.

Bhushan, R., 1989. Firm characteristics and analyst following. Journal of Accounting and Economics 11(2): 255-274.

Bizjak, J., Kalpathy, S., Li, Z.F., and Young, B., 2018. The Role of Peer Firm Selection in Explicit Relative Performance Awards. Working Paper. Available at papers.ssrn.com/sol3/papers.cfm?abstract_id=2833309.

Brennan, M.J. and Kraus, A., 1987. Efficient financing under asymmetric information. The Journal of Finance 42(5): 1225-1243.

Brennan, M.J. and Schwartz, E.S., 1988. The case for convertibles. Journal of Applied Corporate Finance 1(2): 55-64.

Brown, L.D., 1997. Analyst forecasting errors: Additional evidence. Financial Analysts Journal: 81-88.

Brown, L.D., 2001. A temporal analysis of earnings surprises: Profits versus losses. Journal of Accounting Research, 39(2): 221-241.

Brown, L.D. and Caylor, M.L., 2005. A temporal analysis of quarterly earnings thresholds: Propensities and valuation consequences. The Accounting Review 80(2): 423-440.

Celentani, M. and Loveira, R., 2006. A simple explanation of the relative performance evaluation puzzle. Review of Economic Dynamics 9(3): 525-540.

Chang, S.C., Chen, S.S. and Liu, Y., 2004. Why firms use convertibles: A further test of the sequential-financing hypothesis. Journal of Banking & Finance 28(5): 1163-1183.

Chen, Q., Hemmer, T. and Zhang, Y., 2007. On the relation between conservatism in accounting standards and incentives for earnings management. Journal of Accounting Research 45(3): 541-565.

Chen, T., Harford, J. and Lin, C., 2015. Do analysts matter for governance? Evidence from natural experiments. Journal of Financial Economics 115(2): 383-410.

Choi, J.H., Kim, J.B., Qiu, A.A., and Zang, Y., 2012. Geographic proximity between auditor and client: How does it impact audit quality?. Auditing: A Journal of Practice & Theory 31(2): 43-72.

Chung, K.H. and Jo, H., 1996. The impact of security analysts' monitoring and marketing functions on the market value of firms. Journal of Financial and Quantitative analysis 31(4): 493-512.

DeMarzo, P. and Kaniel, R., 2017. Relative Pay for Non-Relative Performance: Keeping up with the Joneses with Optimal Contracts.Working Paper. Available at papers.ssrn.com/sol3/ papers.cfm?abstract_id=2755856.

De Roon, F. and Veld, C., 1998. Announcement effects of convertible bond loans and warrant-bond loans: An empirical analysis for the Dutch market. Journal of Banking & Finance 22(12): 1481-1506.

Dikolli, S.S., Diser, V., Hofmann, C., and Pfeiffer, T., 2018. CEO Power and Relative Performance Evaluation. Contemporary Accounting Research 35(3): 1279-1296.

Dikolli, S.S., Hofmann, C., and Pfeiffer, T., 2013. Relative performance evaluation and peer-performance summarization errors. Review of Accounting Studies 18(1): 34-65.

Dutordoir, M., Lewis, C., Seward, J. and Veld, C., 2014. What we do and do not know about convertible bond financing. Journal of Corporate Finance 24: 3-20.

Dutta, S. and Gigler, F., 2002. The effect of earnings forecasts on earnings management. Journal of Accounting Research 40(3): 631-655.

Dye, R.A., 1992. Relative performance evaluation and project selection. Journal of Accounting Research 30(1): 27-52.

Eckbo, B.E., Masulis, R.W. and Norli, O., 2007. Security offerings. In: Eckbo, B. (Ed.), Handbook of Corporate Finance: Empirical Corporate Finance. Elsevier, Amsterdam, 233–373.

Fershtman, C., Hvide, H., and Weiss, Y., 2003. A Behavioral Explanation of the Relative Performance Evaluation Puzzle. Annals D'économie et de Statistique (71/72): 349-361.

Feriozzi, F., 2011. Paying for observable luck. Rand Journal of Economics 42(2): 387-415.

Fleckinger, P., 2012. Correlation and relative performance evaluation. Journal of Economic Theory 147(1): 93-117.

Francis, J., LaFond, R., Olsson, P.M. and Schipper, K., 2004. Costs of equity and earnings attributes. The Accounting Review 79(4): 967-1010.

Gao, P., 2013. A measurement approach to conservatism and earnings management. Journal of Accounting and Economics 55(2): 251-268.

Garvey, G.T. and Milbourn, T.T., 2006. Asymmetric benchmarking in compensation: Executives are rewarded for good luck but not penalized for bad. Journal of Financial Economics 82(1): 197-225.

Gigler, F., Kanodia, C., Sapra, H., and Venugopalan, R., 2009. Accounting conservatism and the efficiency of debt contracts. Journal of Accounting Research 47(3): 767-797.

Gong, G., Li, L.Y., and Shin, J.Y., 2011. Relative performance evaluation and related peer groups in executive compensation contracts. The Accounting Review 86(3): 1007-1043.

Gopalan, R., Milbourn, T.T., and Song, F., 2010. Strategic Flexibility and the Optimality of Pay for Sector Performance. Review of Financial Studies 23(5): 2060-2098.

Green, R.C., 1984. Investment incentives, debt, and warrants. Journal of Financial Economics 13(1): 115-136.

Gu, Z. and Wu, J.S., 2003. Earnings skewness and analyst forecast bias. Journal of Accounting and Economics 35(1): 5-29.

Hart, O. and Holmstrom, B., 1987. The theory of contracts. Advances in Economic Theory: 294-351.

He, J.J. and Tian, X., 2013. The dark side of analyst coverage: The case of innovation. Journal of Financial Economics 109(3): 856-878.

Herbst, M., 2008. *Oil CEOs: High Prices, Fat Paychecks*. Available at www.bloomberg. com/news/articles/2008-06-17/oil-ceos-high-prices-fat-paychecksbusinessweek-business-news-stock-market-and-financial-advice.

Holmstrom, B., 1979. Moral hazard and observability. The Bell Journal of Economics 10(1): 74-91.

Holmstrom, B., 1982. Moral hazard in teams. The Bell Journal of Economics 13(2): 324-340.

Hong, H. and Kacperczyk, M., 2010. Competition and bias. The Quarterly Journal of Economics 125(4): 1683-1725.

Huang, S.X., Pereira, R. and Wang, C., 2017. Analyst Coverage and the

Likelihood of Meeting or Beating Analyst Earnings Forecasts. Contemporary Accounting Research 34(2): 871-899.

Janakiraman, S.N., Lambert, R.A., and Larcker, D.F., 1992. An empirical investigation of the relative performance evaluation hypothesis. Journal of Accounting Research 30(1): 53-69.

Jongjaroenkamol, P. and Laux, V., 2017. Insider versus outsider CEOs, executive compensation, and accounting manipulation. Journal of Accounting and Economics 63(2): 253-261.

Kahneman, D. and Tversky, A., 1979. Prospect theory: An analysis of decision under risk. Econometrica: Journal of the econometric society 47(2): 263-291.

Kang, J.K. and Stulz, R.M., 1996. How different is Japanese corporate finance? An investigation of the information content of new security issues. Review of Financial Studies 9(1): 109-139.

Kubick, T.R., Lockhart, G.B., Mills, L.F., and Robinson, J.R., 2017. IRS and corporate taxpayer effects of geographic proximity. Journal of Accounting and Economics 63(2-3): 428-453.

Li, J., 2013. Accounting conservatism and debt contracts: Efficient liquidation and covenant renegotiation. Contemporary Accounting Research 30(3): 1082-1098.

Li, X., 2015. Accounting conservatism and the cost of capital: An international analysis. Journal of Business Finance & Accounting 42(5-6): 555-582.

Lim, T., 2001. Rationality and analysts' forecast bias. The Journal of Finance 56(1): 369-385.

Lyandres, E. and Zhdanov, A., 2014. Convertible debt and investment timing. Journal of Corporate Finance 24: 21-37.

Matsumoto, D.A., 2002. Management's incentives to avoid negative earnings surprises. The Accounting Review 77(3): 483-514.

Matsunaga, S.R. and Park, C.W., 2001. The effect of missing a quarterly earnings benchmark on the CEO's annual bonus. The Accounting Review 76(3): 313-332.

Mayers, D., 1998. Why firms issue convertible bonds: the matching of financial and real investment options. Journal of Financial Economics 47(1): 83-102.

Myers, S.C. and Majluf, N.S., 1984. Corporate financing and investment decisions when firms have information that investors do not have. Journal of financial economics 13(2): 187-221.

Olsen B., 2016. As prices plunge, most oil CEOs see drops in pay. Available at http://www. marketwatch.com/story/as-prices-plunge-most-oil-ceos-see-drops-in-pay-2016-04-13.

Penman, S.H. and Zhang, X.J., 2002. Accounting conservatism, the quality of earnings, and stock returns. The Accounting Review 77(2): 237-264.

Previts, G.J., Bricker, R.J., Robinson, T.R., and Young, S.J., 1994. A content analysis of sell-side financial analyst company reports. Accounting

Horizons 8(2): 55-70.

Rahim, N.A., Goodacre, A., and Veld, C., 2014. Wealth effects of convertible-bond and warrant-bond offerings: a meta-analysis. The European Journal of Finance 20(4): 380-398.

Stein, J.C., 1992. Convertible bonds as backdoor equity financing. Journal of Financial Economics 32(1): 3-21.

Venugopalan, R., 2002. Conservatism in accounting: Good or bad? Doctoral dissertation, University of Minnesota.

Weetman, P. and Gray, S.J., 1991. A comparative international analysis of the impact of accounting principles on profits: The USA versus the UK, Sweden and the Netherlands. Accounting and Business Research 21(84): 363-379.

Zhang, X.J., 2000. Conservative accounting and equity valuation. Journal of Accounting and Economics 29(1): 125-149.